NEW DIVISION ALGEBRAS*

BY

L. E. DICKSON

1. **Introduction.** The chief outstanding problem in the theory of linear algebras (or hypercomplex numbers) is the determination of all division algebras. We shall add here very greatly to the present meager knowledge of them, since we shall show how to construct one or more types of division algebras of order n^2 corresponding to every solvable group of order n.

While it was long thought that the theory of continuous groups furnishes an important tool for the study of linear algebras, the reverse position is now taken. But this memoir shows how vital a rôle the theory of finite groups plays in the theory of division algebras.

Fields and the algebra of real quaternions were the only known division algebras until the writer's discovery in 1905 of a division algebra D, over a field F, whose n^2 basal units are $i^a j^b$ $(a, b=0, 1, \cdots, n-1)$, where i is a root of an irreducible cyclic equation of degree n for F. Recently, Cecioni† gave a further division algebra of order 16.

It is here shown that the algebras D form only the first of an infinitude of systems of division algebras. The next system is composed of algebras Γ of order p^2q^2 over F with the basal units $i^a j^b k^c$ $(a<pq, b<q, c<p)$. We start with an irreducible equation of degree pq, three of whose roots are i, and the rational functions $\theta(i)$ and $\psi(i)$ with coefficients in F, such that the qth iterative $\theta^q(i)$ of $\theta(i)$ is i, and likewise $\psi^p(i)=i$, while all the roots are given by

$$\theta^k[\psi^r(i)]=\psi^r[\theta^k(i)] \qquad (k=0,1, \cdots, q-1\ ;\ r=0, 1, \cdots, p-1)\ .$$

The complete multiplication table of the units follows by means of the associative law from

$$i^q=g\ , \qquad k^p=\gamma\ , \qquad kj=\alpha jk\ , \qquad ji=\theta(i)j\ , \qquad ki=\psi(i)k\ ,$$

where g, γ and α are in the field $F(i)$. The conditions for associativity all reduce to

$$g=g(\theta)\ , \qquad \alpha\alpha(\theta)\alpha(\theta^2)\cdots\alpha(\theta^{q-1})g=g(\psi)\ ,$$
$$\gamma=\gamma(\psi)\ , \qquad \alpha\alpha(\psi)\alpha(\psi^2)\cdots\alpha(\psi^{p-1})\gamma(\theta)=\gamma\ .$$

* Presented to the Society, October 31, 1925; received by the editors in October, 1925.

† Rendiconti del Circolo Matematico di Palermo, vol. 47 (1923), pp. 209–54.

The subalgebra with the units $i^a j^b$ may be regarded as an algebra Σ of type D with q^2 units $i^d j^b$ $(d, b=0, 1, \cdots, q-1)$ over the field F_1 derived from F by the adjunction of all the elementary symmetric functions of i, $\theta(i)$, $\theta^2(i), \cdots, \theta^{q-1}(i)$. This Σ is a division algebra if and only if g (which is in F_1) is not the norm, relative to F_1, of any number of the field $F(i)$. For $p=2$ and $p=3$, it is shown that Γ is then a division algebra if and only if

$$\gamma \neq X^{(p-1)} X^{(p-2)} \cdot \cdot \cdot X'' X' X$$

for any X in Σ, where $X^{(r)}$ denotes $k^r X k^{-r}$, which is an element of Σ. This result indicates that we have investigated Γ as a quasi algebra over Σ, where Σ is an algebra over F. Such a treatment is far simpler than the more direct study of Γ as an algebra over F (see the end of §10). The former method accomplishes a factorization of the difficulties both as to the conditions for associativity and the conditions that Γ be a division algebra.

The preceding Γ was obtained from an equation whose Galois group G is an abelian group with two independent generators of orders p and q. There are treated the generalizations when G is an arbitrary abelian group, and when G is any solvable group.

If in the above relations $j^q=g$, etc., we replace each number of the field $F(i)$ by unity, we obtain the generating relations $j^q=1$, $k^p=1$, $kj=jk$ of G. Similarly for any Γ, provided we reverse the order of multiplication (i.e., pass to the reciprocal algebra) in case G is not abelian. While any Γ thus shows at once its G, it is a long story to construct Γ from G.

In §§2, 3 are recast and amplified some known proofs partly to make the paper elementary and self-contained and partly to emphasize the minimum assumptions involved. The paper presupposes only the simplest notions concerning linear algebras* and the elements of the theory of finite groups. Further developments of the paper will be given in Chicago theses.

2. **A class of division algebras.** Let A be any associative division algebra, with the modulus 1, over any field F such that†

(I) A is of order n^2;

(II) A contains an element i which satisfies an equation $f(x)=0$ of degree n irreducible in F;

---

* Cf. the writer's *Algebras and their Arithmetics*, University of Chicago Press, 1923.

† Assumptions I–III do not impose actual limitations on the generality of our study of division algebras A. For, any A may be regarded as an algebra over the field B composed of all those elements of A which are commutative with every element of A. Taking B as a new field F of reference, we call A a *normal* division algebra over F. By *Algebras and their Arithmetics*, pp. 227–28, this normal A is of order a square, say n^2. In the German translation to be published in 1926 by Orell Füssli Verlag, Zürich, I prove at the end of Chapter ~~VIII~~ that A is then of rank n and deduce II and III at once.

XII (p.262-3)

(III) The only elements of A which are commutative with i are polynomials in i with coefficients in F;

(IV) The roots of $f(x)=0$ are all rational functions $\theta_r(i)$ of i with coefficients in F.

All rational functions of i with coefficients in F form a field $F(i)$, whose numbers are known to be expressible as polynomials of degree $<n$ in i with coefficients in F.

There exist elements $z_1=1, z_2, \cdots, z_n$ *of* A *such that every element of* A *can be expressed in one and only one way in the form* $\sum g_k z_k$, *where the* g_k *are in* $F(i)$. For, we may choose as z_2 any element of A not in $F(i)$. If $g_1+g_2z_2$ is equal to $h_1+h_2z_2$, then $g_1=h_1$, $g_2=h_2$. For, their difference $a+bz_2$ is zero. If $b=0$, then $a=0$ and the statement is proved. If $b\neq 0$, z_2 would be the element $-b^{-1}a$ of $F(i)$, contrary to hypothesis. Next, we may choose as z_3 any element not of the form $a+bz_2$, where a and b are in $F(i)$. As before, if $g_1+g_2z_2+g_3z_3$ is equal to $h_1+h_2z_2+h_3z_3$, then $g_k=h_k (k=1, 2, 3)$. In this manner we obtain elements $z_1=1, z_2, \cdots, z_m$ of A such that every element of A can be expressed in one and only one way in the form $\sum g_k z_k$. Hence A is an algebra over F with the nm basal units $i^r z_k$ $(r=0, 1, \cdots, n-1;\ k=1, \cdots, m)$. Since A is of order n^2, we have $nm=n^2$, $m=n$.

Applying this result to the particular element z_si of A, we have

$$z_s i = \sum_{k=1}^{n} g_{sk} z_k \qquad (s=1, \cdots, n)\,, \tag{1}$$

where the g_{sk} are in $F(i)$. Let G be the matrix having the element g_{sk} in the sth row and kth column. Let Z be the matrix composed of a single column of elements z_1 (at the top), $z_2, \cdots, z_n$. Then equations (1) are equivalent to the single equation $Zi=GZ$ in matrices. By induction on r, we get $Zi^r=G^rZ$. Multiply this by the coefficient of i^r in any assigned polynomial $h(i)$ with coefficients in F and sum as to r. We get $Zh(i)=h(G)Z$. Thus $h(G)=0$ implies $z_1h=h(i)=0$. Next, take $h(i)=f(i)=0$. Then $f(G)Z=0$, which implies $f(G)=0$. For, if the element in the sth row and kth column of $f(G)$ is f_{sk}, then $f(G)Z$ is a matrix having a single column whose element in the sth row is $\sum_k f_{sk}z_k=0$. By the remark preceding (1), this implies that every $f_{sk}=0$. The two results show that *the minimum equation* $f(x)=0$ *of* i *for* F *is also the minimum equation of matrix* G *for* F.

Let the minimum equation of G for any field F' containing F be $h(x)=0$ of degree d. Write $f(x)=g(x)h(x)+r(x)$, where $r(x)$ is either identically zero or has a degree $<d$. Then $r(G)=0$, so that $r(x)\equiv 0$. Take F' to be the field $F(i)$ to which belong all elements of G. Thus each root of the minimum equation of G for $F(i)$ is a root of $f(x)=0$. The former has the same roots

apart from multiplicity as the characteristic equation of G (*Algebras*, p. 110). Hence *every root of* $|G-xI|=0$ *is a root of* $f(x)=0$.

Let θ be any root of $|G-xI|=0$. By IV, θ is a rational function $\theta(i)$ of i with coefficients in F. Write $\alpha=\sum a_s z_s$, where the $a_s(s=1,\cdots,n)$ are numbers of $F(i)$. By (1),

$$\alpha i=\sum_{s,k=1}^{n} a_s g_{sk} z_k .$$

Hence $\alpha i=\theta\alpha$ if and only if

$$\sum_{s=1}^{n} g_{sk}a_s-\theta a_k=0 \qquad (k=1,\cdots,n) .$$

The determinant of the coefficients of $a_1,\cdots,a_n$ is $|G-\theta I|$, which is zero. Hence there exist solutions $a_1,\cdots,a_n$, not all zero, which are rational functions of the g_{sk} and $\theta(i)$ with rational coefficients and hence are numbers of $F(i)$. Thus α is in algebra A and $\alpha\neq 0$.

We next prove that the assumption that θ is a multiple root of $|G-xI|=0$ leads to a contradiction. Take the α just found as a new z_1 and retain the notation (1). Then

$$z_1 i=\theta z_1 , \qquad g_{11}=\theta , \qquad g_{1k}=0 \qquad (k>1) .$$

We can find numbers $a_2,\cdots,a_n$, not all zero, and a number c, all in $F(i)$, such that

$$\alpha=\sum_{s=2}^{n} a_s z_s , \qquad \alpha i=\theta\alpha+cz_1 .$$

The conditions are

$$c=\sum_{s=2}^{n} a_s g_{s1} , \qquad \sum_{s=2}^{n} a_s g_{sk}-\theta a_k=0 \quad (k=2,\cdots,n) . \tag{2}$$

Let M_x be the minor of $g_{11}-x$ in $G-xI$. Since the remaining elements $g_{12},\cdots,g_{1n}$ of the first row are zero, $|M_x|$ vanishes when $x=\theta$. The matrix of the coefficients of $a_2,\cdots,a_n$ in equations (2), other than the first, is derived from M_θ by interchanging rows and columns, whence its determinant is zero. Hence those equations can be satisfied by choice of numbers $a_2,\cdots,a_n$, not all zero, of $F(i)$. Thus α is in algebra A and $\alpha\neq 0$. Since α lacks z_1,

$$g(i)z_1+h(i)\alpha=0 \qquad \text{implies } g=h=0 . \tag{2'}$$

If $c=0$, $z_1 i=\theta z_1$ and $\alpha i=\theta\alpha$ imply

$$\alpha^{-1}z_1 i=\alpha^{-1}\theta z_1=i\alpha^{-1}z_1 ,$$

so that $\alpha^{-1}z_1$ is commutative with i and hence by III is a number $\phi(i)$ of $F(i)$. By induction on $k\alpha i,{}^k=\theta^k\alpha$. Multiply by the coefficient of i^k in $\phi(i)$ and sum as to k. Hence

$$\alpha\phi(i)=\phi(\theta)\alpha .$$

Then $z_1 = \alpha\phi(i) = \phi(\theta)\alpha$ contradicts (2′).

Hence $c \neq 0$. Then $\beta = cz_1$ is not zero in the division algebra A, and

$$\beta i = cz_1 i = c\theta z_1 = \theta c z_1 = \theta\beta , \qquad \alpha i = \theta\alpha + \beta .$$

By induction on k,

$$\alpha i^k = \theta^k \alpha + k\theta^{k-1}\beta .$$

Multiply by the coefficient of i^k in $f(i)$ and sum as to k. We get

$$\alpha f(i) = f(\theta)\alpha + f'(\theta)\beta , \qquad 0 = f'(\theta)\beta .$$

This is impossible in a division algebra since $\beta \neq 0$ and $f'(\theta) \neq 0$, there being no double root θ of the irreducible equation $f(x) = 0$. Hence $|G - xI| = 0$ has no multiple root, so that its n distinct roots coincide with the n roots $\theta_r(i)$ of $f(x) = 0$.

We have now proved* that A contains elements $j_r (r = 0, 1, \cdots, n-1)$, each not zero, where $j_0 = 1$, such that

$$j_r i = \theta_r(i) j_r \qquad (r = 0, 1, \cdots, n-1) , \tag{3}$$

$$j_r i^k = [\theta_r(i)]^k j_r , \tag{4}$$

$$j_r \phi(i) = \phi[\theta_r(i)] j_r \qquad (r = 0, 1, \cdots, n-1) . \tag{5}$$

3. **Algebras of types B and C.** Discarding the assumption that A is a division algebra, we shall say that an associative algebra A, having the modulus 1, over a field F, is of *type B* if it has properties I, II, IV, if it contains elements $j_r \neq 0 (r = 1, \cdots, n-1)$ satisfying (3), and finally if every c_{rs} in Lemma 2 is not zero.

LEMMA 1. *A has the basis* $i^k j_r$ $(k, r = 0, 1, \cdots, n-1)$.

For, if these n^2 elements are linearly dependent with respect to F, there exist polynomials $\phi_r(i)$, not all zero, of degree $< n$ in i with coefficients in F such that

$$\sum_{r=0}^{n-1} \phi_r(i) j_r = 0 .$$

Multiply by i^k on the right and apply (4); we get

$$\sum_{r=0}^{n-1} \phi_r(i) [\theta_r(i)]^k j_r = 0 \qquad (k = 0, 1, \cdots, n-1) .$$

The determinant Δ of the coefficients of the $\phi_r(i) j_r$ is equal to the product of the differences of the n distinct roots $\theta_r(i)$ of $f(x) = 0$ and hence is not zero.

* Stated by Wedderburn, these Transactions, vol. 22 (1921), pp. 133–34 (§4). The proof is based on suggestions made by him to the writer.

Multiply* the displayed equation on the left by the cofactor of the element of Δ in the $(k+1)$th row and $(r+1)$th column and sum for k. We get $\Delta\phi_r(i)j_r=0$. If $\phi_r(i)$ is not zero, it has an inverse in $F(i)$, and $j_r=0$, contrary to its origin. Hence every $\phi_r(i)=0$.

Since $f[\theta_s(x)]=0$ is satisfied by the root i of the equation $f(x)=0$ irreducible in F, it is satisfied by the root $\theta_r(i)$ of the latter. Hence $\theta_s[\theta_r(i)]$ is a root $\theta_u(i)$ of $f(x)=0$.

LEMMA 2. *If r and s are any two equal or distinct integers $\leqq n-1$ and if u is the uniquely determined integer for which $\theta_s[\theta_r(i)]=\theta_u(i)$, then*

$$j_r j_s = c_{rs} j_u , \tag{6}$$

where c_{rs} is in $F(i)$.

For by Lemma 1, there exist numbers d_{rst} in $F(i)$ such that

$$j_r j_s = \sum_{t=0}^{n-1} d_{rst} j_t .$$

Multiply by i on the right and apply (3), (5); we get

$$\theta_u(i) j_r j_s = \sum_{t=0}^{n-1} d_{rst}\theta_t(i) j_t .$$

Eliminate $j_r j_s$ and apply Lemma 1. Hence

$$d_{rst}[\theta_u(i)-\theta_t(i)]=0 ,$$

whence $d_{rst}=0$ if $t\neq u$ and therefore $\theta_t(i)\neq\theta_u(i)$.

An algebra of type B shall be called of type C if it has the commutivity property

$$\theta_r[\theta_s(i)]=\theta_s[\theta_r(i)] \qquad (r, s=0, 1, \cdots, n-1) . \tag{7}$$

4. **Algebras of type D.** Consider the case of an algebra of type C for which $f(x)=0$ is an irreducible cyclic equation, having therefore the roots

$$i , \quad \theta_1(i) , \quad \theta_2(i)=\theta_1[\theta_1(i)] , \ \cdots , \ \theta_{r+1}(i)=\theta_1[\theta_r(i)] , \ \cdots ,$$

while $\theta_n(i)=i$. By Lemma 2,

$$j_r j_1 = a_r j_{r+1} \qquad (r=1, \cdots, n-1 \ ; \ j_n=1) .$$

By induction on k, $j_1^k=a_1a_2\cdots a_{k-1}j_k$. Since each $a_r\neq0$ by hypothesis, we may introduce $a_1j_2, a_1a_2j_3, \cdots, a_1\cdots a_{n-2}j_{n-1}$ as new units $j_2, j_3, \cdots, j_{n-1}$, and then have $j_1^k=j_k (k=2, \cdots, n-1)$, while $j_1^n=g$, where g is in $F(i)$. The associative law implies that j_1^k and j_1^n are commutative, whence

* This step is an improvement on the proof by Cecioni, from whom we borrow also Lemma 2.

$gj_k = j_k g = g(\theta_k)j_k$, by (5). Thus $g = g(\theta_k)$ for every k, so that g is a symmetric function of the roots of $f(x)=0$ and hence is in F. Writing j for j_1, we obtain the algebra* D over F having the n^2 basal units $i^s j^k (s,k=0,1,\cdots,n-1)$, where $ji=\theta(i)j, j^n=g$. The condition that g is in F insures that the algebra is associative (§9). It will be shown incidentally in §§12, 13 that for $n=2$ or 3 D is a division algebra if g is not the norm $\Pi\phi[\theta_k(i)]$ of any polynomial $\phi(i)$ in i with coefficients in F. For any n, it was proved by Wedderburn† that D is a division algebra if no power of g less than the nth is the norm of a number of $F(i)$.

5. **Galois group** G. Let G be the Galois group for the field F of the irreducible equation $f(x)=0$ whose roots are rational functions $i_r=\theta_r(i)$ of i with coefficients in F for $r=0,1,\cdots,n-1$. If a substitution of G leaves i unaltered, it leaves each i_r unaltered and is the identity substitution. Since the equation is irreducible, its group G is transitive. Hence G is of order n and contains one and only one substitution Θ_r which replaces i by $\theta_r(i)$. When it is applied to the rational relation $i_l=\theta_l(i)$, we obtain a true relation. Hence Θ_r replaces i_l by $\theta_l[\theta_r(i)]$, which is a certain $i_v=\theta_v(i)$. Similarly, Θ_s replaces i_v by $\theta_v[\theta_s(i)]$. Hence the product $\Theta_r\Theta_s$ replaces i_l by

$$\theta_v[\theta_s i] = \theta_l[\theta_r(\theta_s i)] = \theta_l[\theta_t(i)]\ , \quad \text{if} \quad \theta_r[\theta_s(i)] = \theta_t(i)\ .$$

But Θ_t replaces i_l by $\theta_l[\theta_t(i)]$. Hence

$$\Theta_r\Theta_s = \Theta_t \quad \text{if and only if} \quad \theta_r[\theta_s(i)] = \theta_t(i)\ . \tag{8}$$

First, let $f(x)=0$ be an abelian equation so that the roots have the commutivity property (7). Then $\Theta_r\Theta_s=\Theta_s\Theta_r$ by (8), and G is an abelian (commutative) group. It is well known that any abelian group G has a set of independent generators $g_1,\cdots,g_k$ such that every substitution of G can be expressed in one and only one way in the form $g_1^{e_1}\cdots g_k^{e_k}$ $(e_i=0,1,\cdots,h_i-1)$, where h_i is the order of g_i. Select any one of the independent generators of G, denote its order by p, and write $q=n/p$. Adopting a new subscript notation for the n substitutions of G, we shall write Θ_q for the selected generator and write $\Theta_0=1, \Theta_1,\cdots,\Theta_{q-1}$ for the subsitutions of the subgroup G_q which is generated by the generators other than Θ_q of G. The substitutions of G are therefore

$$\Theta_q^r\Theta_k \qquad (r=0,1,\cdots,p-1\ ;\ k=0,1,\cdots,q-1)\ .$$

---

* Discovered by the writer and announced as a division algebra in the Bulletin of the American Mathematical Society, vol. 22 (1905–06), p. 442; details in these Transactions, vol. 15 (1914), p. 31.

† These Transactions, vol. 15 (1914), p. 162. Amplified in Dickson's *Algebras and their Arithmetics*, 1923, p. 221.

Since we have defined Θ_s only when $s \leqq q$, we are at liberty to write

$$\Theta_{k+rq}=\Theta_q^r\Theta_k \ .$$

By (8) we see that the roots have now been assigned a subscript notation such that $\theta_q^p(i)=i$ and

$$(9) \qquad \theta_{rq}(i)=\theta_q^r(i) \ , \qquad \theta_{k+rq}(i)=\theta_{rq}[\theta_k(i)] \qquad (r<p \ , \ k<q) \ ,$$

where $\theta^r(i)$ is the rth iterative of the function $\theta(i)$.

Second, let G be not abelian. We assume that G has an invariant subgroup G_q which is extended to G by a substitution Θ_q (whose order may exceed the index p of G_q under G). Since Θ_q transforms each substitution Θ_k of G_q into a substitution Θ_{k_0} of G_q, we have $\Theta_k\Theta_q=\Theta_q\Theta_{k_0}$. Hence in any product of factors $\Theta_k(k<q)$ and Θ_q, we can move the factors Θ_q to the front. Thus every substitution of G can be expressed in the form $\Theta_q^r\Theta_k(k<q)$. If the sth, but no lower, power of Θ_q belongs to G_q, the preceding formula with $r=0, 1, \cdots, s-1$; $k=0, 1, \cdots, q-1$, gives all the pq substitutions of G without repetition, whence $s=p$. Hence we may assign a subscript notation to the roots such that (9) holds. We assume that G_q has an invariant subgroup $G_{q'}$ which is extended to G_q by a single substitution; similarly for $G_{q''}$, etc. These assumptions imply that G is a solvable group. For, let p be the product of the primes $a_1, \cdots, a_e$, not necessarily distinct. Then G_q is an invariant subgroup of index a_1 of the extension G_{a_1q} of G_q by Θ_q^{p/a_1}, the a_1th, but no lower, power of which is in G_q. Similarly, G_{a_1q} is an invariant subgroup of index a_2 of the extension $G_{a_1a_2q}$ of G_q by $\Theta_q^{p/(a_1a_2)}$, whose a_2th power is the former extender. Finally, if $\pi=a_1a_2\cdots a_{e-1}$, $G_{\pi q}$ is an invariant subgroup of index a_e of $G=G_{pq}$. The same argument applies to $G_{q'}$, etc. Hence we may proceed from G to the identity group through a series of groups each an invariant subgroup of prime index of its predecessor. This is the definition of a solvable group.

Conversely, any solvable group serves as a G. For, it has an invariant subgroup G_q of prime index a. We may take $p=a$ and Θ_q to be any substitution of G which is not in G_q. If the sth, but no lower, power of Θ_q belongs to G_q, we see as above that Θ_q extends G_q to a group of order sq, which must divide pq, whence $s=p$.

Algebras which satisfy the present and earlier assumptions shall be said to be of type E.

6. **Algebra** Σ. Let Γ be an algebra over F of type E. We saw that the Θ_r $(0\leqq r<q)$ form a group G_q. Hence if r and s belong to the set $0, 1, \cdots, q-1$, we can find an integer u of the same set such that $\Theta_s\Theta_r=\Theta_u$. By (8), we have

(6), whence the totality of linear functions of $1, j_1, \cdots, j_{q-1}$ with coefficients in $F(i)$ is a subalgebra Σ of Γ. Thus Σ is of order $nq=pq^2$ over F.

Each $j_s (s<q)$ is commutative with every elementary symmetric function E of $\theta_0(i)=i, \cdots, \theta_{q-1}(i)$. For by (5), $j_s E = Hj_s$, where H is the same elementary symmetric function of $\theta_0[\theta_s(i)], \cdots, \theta_{q-1}[\theta_s(i)]$, which by (8) are equal to $\theta_0(i), \cdots, \theta_{q-1}(i)$ in a new order.

Let F_1 be the field obtained by adjoining to F all these functions E. Then i is a root of the equation

$$f_1(x)=\prod_{k=0}^{q-1}[x-\theta_k(i)]=0 ,$$

whose coefficients belong to F_1. *This equation is irreducible in* F_1. For, suppose that $f_1(x)$ has a factor $q(x)$ with coefficients in F_1 which vanishes for $x=i$, but not for $\theta_s(i)$. Let e be any elementary symmetric function of the roots of $q(x)=0$. Thus $\pm e$ is equal to a coefficient and belongs to F_1. Hence e is equal to a polynomial in the adjoined E's and hence is a symmetric function of $\theta_0(i), \cdots, \theta_{q-1}(i)$ with coefficients in F. Thus e is commutative with every $j_k (k<q)$. Also $j_s e = hj_s$, where h is the same elementary symmetric function of the $\theta_k[\theta_s(i)]$, where the $\theta_k(i)$ are the roots of $q(x)=0$. Thus $ej_s = hj_s$, whence $e=h$. Since this is true for every e, the $\theta_k(i)$ coincide in some order with the $\theta_k[\theta_s(i)]$. But for $k=0$ the latter is $\theta_s(i)$, which is not one of the roots $\theta_k(i)$ of $q(x)=0$.

THEOREM 1. *Algebra Σ of order pq^2 over F may be regarded as an algebra of type E of order q^2 over F_1 obtained by means of an equation $f_1(x)=0$ of degree q which is irreducible in F_1. Here F_1 is the field derived from F by adjoining the elementary symmetric functions of the roots $\theta_0(i)=i, \cdots, \theta_{q-1}(i)$ of $f_1(x)=0$.*

ALGEBRAS WITH AN ABELIAN GROUP G, §§7–14

7. **Γ as an algebra over Σ.** By (9) and Lemma 2,

$$j_q^2=c_2 j_{2q}, \; j_q^3=c_3 j_{3q}, \; \cdots, \; j_q^{p-1}=c_{p-1} j_{(p-1)q} ,$$

where each c_k is a number $\neq 0$ of $F(i)$. The second members may be introduced as new units $j_{2q}, j_{3q}, \cdots$. Then

$$j_q^r=j_{rq} \quad (r=2, \cdots, p-1), \; j_q^p=\gamma\neq 0 , \tag{10}$$

where γ is in $F(i)$. In the same manner,

$$j_k j_{rq}=d_{kr} j_{k+rq} \qquad (k=1, \cdots, q-1 \; ; r=1, \cdots, p-1) ,$$

where each d_{kr} is a number $\neq 0$ of $F(i)$, and the second members may be introduced as new units j. Thus

(11) $j_k j_{rq} = j_{k+rq}$, $\quad j_q j_k = \alpha_k j_{q+k} \quad (k=1, \cdots, q-1\,;\, r=1, \cdots, p-1)$,

where each α_k is a number $\neq 0$ of $F(i)$. By the first relations in (10) and (11), any $j_s (s \geqq q)$ may be expressed as a product of a certain $j_k (0 \leqq k < q)$ by a certain j_q^r. Hence every element of algebra Γ is of the form

(12) $$\mathscr{A} = \sum_{k=0}^{p-1} A_k j_q^k\,,$$

where each A_k is in Σ, being of the form

(13) $$A = \sum_{k=0}^{q-1} f_k j_k\,,$$

where f_k is a polynomial $f_k(i)$ in i with coefficients in F.

By (5) and (11_2),

$$j_q A = f_0(\theta_q) j_q + \sum_{k=1}^{q-1} f_k(\theta_q) \alpha_k j_{q+k}\,.$$

The final j is equal to $j_k j_q$ by (11_1). Hence

(14) $$j_q A = A' j_q\,, \qquad A' = f_0(\theta_q) + \sum_{k=1}^{q-1} f_k(\theta_q) \alpha_k j_k \quad \text{for } A \text{ in (13)}\,.$$

Thus j_k transforms any element A of Σ into an element of A' of Σ. Write A'' for $(A')'$, $\cdots$, $A^{(r+1)}$ for $(A^{(r)})'$. By induction,

(15) $$j_q^r A = A^{(r)} j_q^r\,.$$

In particular, since $\gamma = j_q^p$ is commutative with j_q,

(16) $$\gamma' = \gamma \neq 0\,.$$

By (15) for $r = p+s$ and (10_2),

$$A^{(p+s)} \gamma j_q^s = \gamma j_q^s A = \gamma A^{(s)} j_q^s\,,$$

(17) $$A^{(p+s)} = \gamma A^{(s)} \gamma^{-1}\,.$$

For any elements A and B in Σ,

$$(AB)^{(r)} j_q^r = j_q^r AB = A^{(r)} j_q^r B = A^{(r)} B^{(r)} j_q^r\,,$$

(18) $$(AB)^{(r)} = A^{(r)} B^{(r)}\,, \qquad (A+B)^{(r)} = A^{(r)} + B^{(r)}\,.$$

Let

(19) $$\mathscr{B}=\sum_{l=0}^{p-1} B_l j_q^l\ .$$

Then

(20) $$\mathscr{A}\mathscr{B}=\mathscr{P}=\sum_{s=0}^{p-1} P_s j_q^s\ ,$$
$$P_s=\sum_{k=0}^{s} A_k B_{s-k}^{(k)}+\sum_{k=s+1}^{p-1} A_k B_{p+s-k}^{(k)}\gamma\ .$$

This gives the product of any two elements of Γ as an element of Γ.

Heretofore we have assumed that Γ is associative and deduced various needed formulas. We shall now proceed conversely and abstractly and establish the following result.

THEOREM 2. *Let Σ be an associative algebra to every element A of which corresponds a unique element A' of Σ. Define $A''=(A')', \cdots, A^{(r)}=(A^{(r-1)})'$, so that $(A^{(r)})^{(s)}=A^{(r+s)}$. Let Σ contain an element $\gamma=\gamma'\neq 0$. Let the $A^{(r)}$ have the properties* (17) *and* (18). *Consider the set Γ of elements $\mathscr{A}=(A_0, A_1, \cdots, A_{p-1})$ in which the A_k range independently over Σ. Define the product of $\mathscr{A}$ by $\mathscr{B}=(B_0, \cdots, B_{p-1})$ to be $\mathscr{P}=(P_0, \cdots, P_{p-1})$, where P_s is given by* (20). *Then this multiplication is associative.*

We shall now give a direct proof, and later an indirect proof (§8). Let $\mathscr{C}=(C_0, \cdots, C_{p-1})$ be any third element of Γ. Then $\mathscr{P}\mathscr{C}=\mathscr{D}$, where

$$D_r=\sum_{s=0}^{r} P_s C_{r-s}^{(s)}+\sum_{s=r+1}^{p-1} P_s C_{p+r-s}^{(s)}\gamma\ .$$

Inserting the value (20) of P_s, we get $D_r=d_1+d_2+d_3+d_4$, where

$$d_1=\sum_{s=0}^{r}\sum_{k=0}^{s} A_k B_{s-k}^{(k)} C_{r-s}^{(s)}\ ,\qquad d_2=\sum_{s=0}^{r}\sum_{k=s+1}^{p-1} A_k B_{p+s-k}^{(k)}\gamma C_{r-s}^{(s)}\ ,$$
$$d_3=\sum_{s=r+1}^{p-1}\sum_{k=0}^{s} A_k B_{s-k}^{(k)} C_{p+r-s}^{(s)}\gamma\ ,\quad d_4=\sum_{s=r+1}^{p-1}\sum_{k=s+1}^{p-1} A_k B_{p+s-k}^{(k)}\gamma C_{p-r-s}^{(s)}\gamma\ .$$

Next, $\mathscr{B}\mathscr{C}=\mathfrak{Q}=(Q_0, \cdots)$, where

$$Q_s=\sum_{t=0}^{s} B_t C_{s-t}^{(t)}+\sum_{t=s+1}^{p-1} B_t C_{p+s-t}^{(t)}\gamma\ ,$$

and $\mathscr{A}\mathfrak{Q}=\mathscr{E}=(E_0, \cdots)$, where

$$E_r=\sum_{k=0}^{r} A_k Q_{r-k}^{(k)}+\sum_{k=r+1}^{p-1} A_k Q_{p+r-k}^{(k)}\gamma\ .$$

We insert the value of $Q_s^{(k)}$ found from Q_s by use of (18) and $\gamma^{(k)}=\gamma$. We get $E_r=M+R+S+T$, where

$$M=\sum_{k=0}^{r}\sum_{t=0}^{r-k}A_kB_t^{(k)}C_{r-k-t}^{(t+k)}\,,\qquad R=\sum_{k=0}^{r}\sum_{t=r-k+1}^{p-1}A_kB_t^{(k)}C_{p-t+r-k}^{(t+k)}\gamma\,,$$

$$S=\sum_{k=r+1}^{p-1}\sum_{t=0}^{p+r-k}A_kB_t^{(k)}C_{p+r-k-t}^{(t+k)}\gamma\,,\qquad T=\sum_{k=r+1}^{p-1}\sum_{t=p+r-k+1}^{p-1}A_kB_t^{(k)}C_{2p-t+r-k}^{(t+k)}\gamma\gamma\,.$$

In M write $s=t+k$; we get d_1. In T, $t+k\geqq p+r+1>p$; we write $s=t+k-p$, apply (17), and get

$$T=\sum_{k=r+1}^{p-1}\sum_{s=r+1}^{k-1}A_kB_{s-k+p}^{(k)}\gamma C_{p+r-s}^{(s)}\gamma=\sum_{s=r+1}^{p-2}\sum_{k=s+1}^{p-1}=d_4\,.$$

In the terms of R having $t+k<p$, we set $s=t+k$ and get R_1; in those having $t+k\geqq p$, we set $s=t+k-p$, apply (17), and get R_2:

$$R_1=\sum_{k=0}^{r}\sum_{s=r+1}^{p-1}A_kB_{s-k}^{(k)}C_{p+r-s}^{(s)}\gamma\,,\qquad R_2=\sum_{k=0}^{r}\sum_{s=0}^{k-1}A_kB_{p+s-k}^{(k)}\gamma C_{r-s}^{(s)}=\sum_{s=0}^{r}\sum_{k=s+1}^{r}\,.$$

Treating S as we did R, we get $S=S_1+S_2$,

$$S_1=\sum_{k=r+1}^{p-1}\sum_{s=k}^{p-1}A_kB_{s-k}^{(k)}C_{p+r-s}^{(s)}\gamma=\sum_{s=r+1}^{p-1}\sum_{k=r+1}^{s}\,,$$

$$S_2=\sum_{k=r+1}^{p-1}\sum_{s=0}^{r}A_kB_{p+s-k}^{(k)}\gamma C_{r-s}^{(s)}\,.$$

In R_1 and S_2 we may interchange the summation signs since the limits are constants. We see that $R_1+S_1=d_3$, $R_2+S_2=d_2$. Hence $\mathscr{AB}\cdot\mathscr{C}=\mathscr{A}\cdot\mathscr{BC}$ for any three elements of Γ.

8. **Γ as an algebra of matrices with elements in Σ.** To $\mathscr{A}$ in (12) we make correspond the square matrix

$$(21)\quad\left(\begin{array}{ccccccc}
A_0 & A_1 & A_2 & \cdots & A_{p-3} & A_{p-2} & A_{p-1}\\
A'_{p-1}\gamma & A'_0 & A'_1 & \cdots & A'_{p-4} & A'_{p-3} & A'_{p-2}\\
A''_{p-2}\gamma & A''_{p-1}\gamma & A''_0 & \cdots & A''_{p-5} & A''_{p-4} & A''_{p-3}\\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots\\
A_2^{(p-2)}\gamma & A_3^{(p-2)}\gamma & A_4^{(p-2)}\gamma & \cdots & A_{p-1}^{(p-2)}\gamma & A_0^{(p-2)} & A_1^{(p-2)}\\
A_1^{(p-1)}\gamma & A_2^{(p-1)}\gamma & A_3^{(p-1)}\gamma & \cdots & A_{p-2}^{(p-1)}\gamma & A_{p-1}^{(p-1)}\gamma & A_0^{(p-1)}
\end{array}\right),$$

denoted by $[\mathscr{A}]$, in which any row is derived from the preceding row by permuting its elements cyclically, adding another accent, and multiplying the new first element by $\gamma=\gamma'$ on the right. Using $\gamma'=\gamma$, (18) and the case $A^{(p)}=\gamma A\gamma^{-1}$ of (17), we find that* $[\mathscr{A}][\mathscr{B}]=[\mathscr{P}]$, where the P_s are given by (20). However, it is sufficient to compute only the elements of the first row of the product in view of the corollary below.

THEOREM 3. *Let $M=(m_{ij})$ be any p-rowed square matrix and $M'=(m'_{ij})$. Let T be the p-rowed square matrix† whose elements one place to the right of the diagonal are all* 1, *whose first element in the last row is $\gamma=\gamma'\neq 0$, and whose further elements are all zero. Then $TMT^{-1}=M'$ if and only if M is of the form* (21) *and $A_i^{(p)}=\gamma A_i\gamma^{-1}(i=0,1,\cdots,p-1)$, where $A''=(A')',\cdots,A^{(k)}=A'^{(k-1)}$.*

For, if we multiply the first row of M on the left by γ and carry it to the new last row, we get TM. If we multiply the last column of M' by γ on the right and carry it to the new first column, we get $M'T$. Thus $TM=M'T$ if and only if the second row of M (which is the first row of TM) is the first row of $M'T$ and hence is derived from the first row of M' by permuting its elements cyclically and multiplying the new first element by γ on the right, and if the third row of M (second of TM) is derived as before from the second row of M' and hence is derived in the same way from the second row of M with the addition of another accent, . . . , and finally if we permute cyclically the elements of the last row of M, add an accent, and multiply the new first element by γ on the right we get the last row of TM (which is the product of the first row of M by γ on the left).

Let also a second matrix N have the property that $TNT^{-1}=N'$. For any two elements m and n of M and N, let $(mn)'=m'n'$. Then evidently $(MN)'=M'N'$. Thus

$$TMNT^{-1}=M'N'=(MN)' .$$

Hence by Theorem 3, MN is of the form (21).

COROLLARY. *Under the assumptions* (18), $\gamma'=\gamma\neq 0$, *and* $A^{(p)}=\gamma A\gamma^{-1}$, *the product of any two matrices of type* (21) *is of that type.*

9. **Associativity conditions.** We seek the conditions on the constants of multiplication of Γ that Γ be an associative algebra. For A' defined in (14),

---

* Under the usual rule of multiplication of matrices, provided an element of $[\mathscr{A}]$ is kept as a left factor and one of $[\mathscr{B}]$ as a right factor.

† For $p=3$, $T=\begin{pmatrix}0&1&0\\0&0&1\\\gamma&0&0\end{pmatrix}$.

we see that $(A+B)'=A'+B'$ is satisfied identically. Hence $(AB)'=A'B'$ for all A and B in Σ if and only if it holds for $A=fj_k$, $B=hj_r (k, r=0, 1, \cdots, q-1)$, where f and h are arbitrary numbers of $F(i)$. This property holds identically if $k=0$ or $r=0$. Hence let $k>0$, $r>0$. Then

$$A'=f(\theta_q)\alpha_k j_k\,, \qquad B'=h(\theta_q)\alpha_r j_r\,.$$

By Lemma 2, $j_k j_r = c_{kr} j_u$, where c_{kr} is in $F(i)$ and u is determined so that $\theta_r[\theta_k(i)]=\theta_u(i)$. Thus

$$A'B'=f(\theta_q)\alpha_k h[\theta_q(\theta_k i)]\alpha_r(\theta_k)c_{kr}j_u\,,$$
$$(AB)'=f(\theta_q)h[\theta_k(\theta_q i)]c_{kr}(\theta_q)\alpha_u j_u\,,$$

if we take $\alpha_0=1$. The two h's are equal by (7). Hence*

$$\alpha_k\alpha_r(\theta_k)c_{kr}=c_{kr}(\theta_q)\alpha_u \qquad (k, r=1, \cdots, q-1;\ \alpha_0=1). \tag{22}$$

It remains to consider $A^{(p)}=\gamma A\gamma^{-1}$. From (14) we get by induction

$$A^{(s)}=f_0(\theta_q^s)+\sum_{k=1}^{q-1} f_k(\theta_q^s)\alpha_k(\theta_q^{s-1})\alpha_k(\theta_q^{s-2})\cdots\alpha_k(\theta_q)\alpha_k j_k\,,$$

where θ_q^s denotes the sth iterative of $\theta_q(i)$. By (9), $\theta_q^p=i$. Taking $s=p$, we get

$$\alpha_k\alpha_k(\theta_q)\alpha_k(\theta_q^2)\cdots\alpha_k(\theta_q^{p-1})=\gamma\gamma^{-1}(\theta_k) \quad (k=1, \cdots, q-1). \tag{23}$$

Finally, $\gamma=\gamma'$ if and only if

$$\gamma=\gamma(\theta_q)\,. \tag{24}$$

THEOREM 4. *If Σ is associative, algebra Γ with an abelian group is associative if and only if conditions* (22)–(24) *hold.*

If $q=1$, conditions (22) and (23) are satisfied vacuously, while Σ is the field $F(i)$ and Γ is of type D of §4.

COROLLARY. *An algebra of type D is associative if and only if $\gamma=\gamma(\theta_1)$, which implies that γ is in F.*

10. **Algebras Γ whose abelian group has two generators.** The subgroup G_q is now cyclic. Hence by Theorem 1, algebra Σ is now of type D of order q^2 over the field F_1 derived from F by adjoining all the symmetric functions of $i, \theta_1, \cdots, \theta_{q-1}$. By §4, we may take $j_k=j_1^k (k=2, \cdots, q-1)$, $j_1^q=g$, where g is in F_1, a necessary and sufficient condition for which is $g=g(\theta_1)$. In (6),

* Note that (22) is the condition that we obtain equal results when we express $j_q\cdot j_k j_r$ and $j_q j_k\cdot j^r$ in the form $(\ )j_q$. This interpretation is useful when we consider a Γ whose G has more than two generators.

$$j_r j_k = j_u\,,\quad u=r+k\,,\quad c_{rk}=1 \qquad (r+k<q)\,,$$

$$j_r j_k = g j_u\,,\quad u=r+k-q\,,\quad c_{rk}=g \qquad (r+k\geqq q\,,\ r<q\,,\ k<q)\,.$$

Hence associativity conditions (22) become

$$\alpha_k \alpha_r(\theta_1^k) = \alpha_{r+k} \qquad (r,\ k=1,\cdots,q-1\ ;\ r+k<q)\,, \tag{25}$$

$$\alpha_k \alpha_r(\theta_1^k) g = g(\theta_q)\alpha_{r+k-q} \qquad (r<q\,,\ k<q\,,\ r+k\geqq q)\,, \tag{26}$$

where $\alpha_0=1$. Write α for α_1. For $r=1$, (25) gives by induction

$$\alpha_k = \alpha\alpha(\theta_1)\alpha(\theta_1^2)\cdots\alpha(\theta_1^{k-1}) \qquad (k=1,\cdots,q-1)\,. \tag{27}$$

Then every (25) follows at once from (27). For $r+k=q$ in (26), we replace the α's by their values (27), note that the final α is $\alpha_0=1$, and get at once

$$\alpha\alpha(\theta_1)\alpha(\theta_1^2)\cdots\alpha(\theta_1^{q-1})g = g(\theta_q)\,. \tag{28}$$

For $r+k>q$, the value from (27) of the final α in (26) cancels with part of the the left member, whence

$$\alpha(\theta_1^{r+k-q})\alpha(\theta_1^{r+k-q+1})\cdots\alpha(\theta_1^{r+k-1})g = g(\theta_q)\,.$$

The q exponents form an arithmetical progression with the common difference 1 and hence are congruent modulo q to $0, 1, \cdots, q-1$ in some order. Hence the relation is (28).

Relations (23) all follow from the case $k=1$:

$$\alpha\alpha(\theta_q)\alpha(\theta_q^2)\cdots\alpha(\theta_q^{p-1}) = \gamma\gamma^{-1}(\theta_1)\,. \tag{29}$$

For, if in (29) we replace i by $i, \theta_1, \theta_1^2, \cdots, \theta_1^{k-1}$ in turn, multiply together the resulting equations and apply (27), we obtain (23). Hence Γ is associative if and only if $g=g(\theta_1)$, $\gamma=\gamma(\theta_q)$, and (27), (28), (29) hold. Of these, (27) serve merely to express the α_k in terms of α.

THEOREM 5. *Let $f(x)=0$ be an equation of degree pq irreducible in a field F whose Galois group for F is abelian and has two independent generators Θ_1 and Θ_q of orders q and p respectively. Then its roots* are*

$$\theta_1^k[\theta_q^r(i)] = \theta_q^r[\theta_1^k(i)] \qquad (k=0,1,\cdots,q-1\ ;\ r=0,1,\cdots,p-1)\,,$$

where θ_1 and θ_q are rational functions of i with coefficients in F such that the

* We may ignore the group and start with any irreducible equation whose pq roots are of the specified type.

qth iterative $\theta_1^q(i)$ of $\theta_1(i)$ is i and likewise $\theta_q^p(i)=i$. There exists an associative algebra Σ whose elements are

$$f_0+f_1j_1+f_2j_1^2+\cdots+f_{q-1}j_1^{q-1}, \tag{30}$$

where the f_k are polynomials in i of degree $\leqq pq-1$ with coefficients in F, while

$$j_1^q=g(i)=g(\theta_1)\ , \qquad j_1^r\phi(i)=\phi[\theta_1^r(i)]j_1^r \qquad (r=1,\cdots,q-1)\ ,$$

so that the product of any two elements (30) *of Σ is another element* (30) *of Σ. Under multiplication defined by* (20), *the totality of polynomials in j_q with coefficients in Σ form an algebra Γ of order p^2q^2 over F which is associative if and only if the four conditions* (24), (28),(29), *and $g=g(\theta_1)$ hold between the parameters g,γ,α of Γ.*

The conditions for associativity are not inconsistent since they are satisfied when $\alpha=1$ and g and γ are both in F. In this case the constants of multiplication of Γ all belong to F. It follows by induction from (22)–(24) that the corresponding results hold when G is abelian and has any number of generators.

Note that Γ has the basis $i^rj_1^aj_q^b$ $(r<pq,\ a<q,\ b<p)$. The laws of multiplication follow readily by the associative law from

$$j_1^q=g\ , \quad j_q^p=\gamma\ , \quad j_qj_1=\alpha j_1j_q\ , \quad j_ti=\theta_t(i)j_t \qquad (t=1,q)\ . \tag{31}$$

We find by induction that

$$j_q^rj_1^k=A_{rk}j_1^kj_q^r\ , \quad A_{rk}=\alpha_k\alpha_k(\theta_q)\cdots\alpha_k(\theta_q^{r-1})\ , \tag{32}$$

while $\alpha_k=\alpha\alpha(\theta_1)\cdots\alpha(\theta_1^{k-1})$, in accord with (27). Then for any polynomials u and w in $F(i)$, we have

$$uj_1^aj_q^b\cdot wj_1^cj_q^d=uw(\theta_1^a\theta_q^b)A_{bc}(\theta_1^a)j_1^{a+c}j_q^{b+d}\ . \tag{33}$$

If $zj_1^fj_q^h$ is any third element of Γ, we find that the associative law holds if and only if*

$$A_{bc}A_{b+d,f}(\theta_1^c)=A_{b,c+f}A_{df}(\theta_1^c\theta_q^b)\ , \tag{34}$$

which is independent of a,h,u,w,z. Since (34) involves four indices b,c,d,f, whereas (25), etc., involved only two, the present direct method of finding the conditions for associativity is far more complicated than our earlier indirect method.

11. **Representation as a matric algebra.** We shall represent algebra

* The initial condition was (34) with i replaced by θ^a If $f=1$, (34) reduces to (25).

Γ of §10 as an algebra of matrices with elements in $F(i)$. First we give such a representation of its subalgebra Σ of elements (30). We make use of §8 with Γ and Σ replaced by our present Σ and $F(i)$, respectively, and $j_q A = A' j_q$ in (14) replaced by $j_1 a = a(\theta_1) j_1$, where a is in $F(i)$. Hence to the element f given by (30) we make correspond the matrix

$$(35)\quad \left\{\begin{array}{ccccc} f_0 & f_1 & f_2 & \cdots & f_{q-1} \\ f_{q-1}(\theta_1)g & f_0(\theta_1) & f_1(\theta_1) & \cdots & f_{q-2}(\theta_1) \\ f_{q-2}(\theta_1^2)g & f_{q-1}(\theta_1^2)g & f_0(\theta_1^2) & \cdots & f_{q-3}(\theta_1^2) \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ f_1(\theta_1^{q-1})g & f_2(\theta_1^{q-1})g & f_3(\theta_1^{q-1})g & \cdots & f_0(\theta_1^{q-1}) \end{array}\right\},$$

which we denote by (f). Since $g(\theta_1)=g$, and $\theta_1^q(i)=i$, it follows from §8 without further discussion that $fr=s$ implies $(f)(r)=(s)$. These results have been established otherwise,* since Σ is an algebra of type D of order q^2 over F_1.

In matrix (21) replace each entry f, which is an element of Σ, by its matric representation (f). Such a p-rowed matrix $[\mathscr{A}]$ whose elements are q-rowed matrices gives rise at once to a pq-rowed matrix $\{\mathscr{A}\}$, formed by erasing the parentheses which enclose the elements of the sub-matrices. For example, let $p=q=2$, $A_0=a+bj_1$, $A_1=c+dj_1$, whence $A_0'=a_2+b_2\alpha j_1$, where a_2 denotes $a(\theta_2)$. Writing a_k for $a(\theta_k)$, $\theta_3(i)=\theta_1[\theta_2(i)]$, we have from (35)

$$(A_0)=\begin{pmatrix} a & b \\ b_1 g & a_1 \end{pmatrix}, \qquad (A_1)=\begin{pmatrix} c & d \\ d_1 g & c_1 \end{pmatrix},$$

and then from (21)

$$(36)\qquad \{\mathscr{A}\}=\left\{\begin{array}{cccc} a & b & c & d \\ b_1 g & a_1 & d_1 g & c_1 \\ c_2\gamma & d_2\alpha\gamma_1 & a_2 & b_2\alpha \\ d_3\alpha_1\gamma g & c_3\gamma_1 & b_3\alpha_1 g & a_3 \end{array}\right\},$$

since $A_1'\gamma=c_2\gamma+d_2\alpha j_1\gamma$, $j_1\gamma=\gamma_1 j_1$, $\gamma_2=\gamma$.

The matrices $\{\mathscr{A}\}$ give the desired matric representation of Γ since the equation $\mathscr{A}\mathscr{B}=\mathscr{P}$ between any three elements of Γ implies $[\mathscr{A}][\mathscr{B}]=[\mathscr{P}]$, by §8, which in turn implies $\{\mathscr{A}\}\{\mathscr{B}\}=\{\mathscr{P}\}$. This is due to the following general theorem. Let (a) be a pq-rowed square matrix having a_{ij} as the element in the ith row and jth column. Let $(a)(b)=(c)$. By grouping the

* *Algebras and their Arithmetics*, pp. 221–24.

successive rows of (a) in blocks of q each and its successive columns in blocks of q, we obtain a compound matrix

$$[A]=\begin{pmatrix} A_{q.q} & A_{q.2q} & \cdots & A_{q.pq} \\ \cdot & \cdot & \cdots & \cdot \\ A_{pq.q} & A_{pq.2q} & \cdots & A_{pq.pq} \end{pmatrix},$$

where $A_{r.s}$ is a q-rowed square matrix having a_{rs} as the last element of the last row. For example, if $p=q=2$,

$$A_{22}=\begin{pmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{pmatrix}, \qquad A_{24}=\begin{pmatrix} a_{13} & a_{14} \\ a_{23} & a_{24} \end{pmatrix}, \qquad A_{42}=\begin{pmatrix} a_{31} & a_{32} \\ a_{41} & a_{42} \end{pmatrix},$$

$$A_{44}=\begin{pmatrix} a_{33} & a_{34} \\ a_{43} & a_{44} \end{pmatrix}.$$

The element in the ith row and jth column of $[A][B]$ is

$$\sum_{s=1}^{p} A_{iq.sq}B_{sq,jq}=C_{iq,jq} ,$$

whence $[A][B]=[C]$. Conversely, this implies $(a)(b)=(c)$.

The present method evidently leads to a representation by matrices with elements in $F(i)$ of any algebra Γ whatever be the number of generators of its abelian group. The existence of such a representation was proved by Cecioni by means of various technical theorems on linear algebras and matrices.

12. **Division algebras Γ with $p=2$.** We assume that Σ is a division algebra. Multiplication (20) in Γ becomes for $p=2$

$$(A_0+A_1j_q)(B_0+B_1j_q)=(A_0B_0+A_1B_1'\gamma)+(A_0B_1+A_1B_0')j_q . \tag{37}$$

Let this product be zero while neither factor is zero. If $B_1=0$, then $B_0\neq 0$, $B_0'\neq 0$, $A_0B_0=0$, $A_1B_0'=0$, $A_0=A_1=0$, contrary to hypothesis. Since $B_1\neq 0$ it has an inverse in Σ, and the product is equal to

$$[(A_0+A_1j_q)B_1][B_1^{-1}B_0+j_q] .$$

Hence it suffices to treat the case $B_1=1$. Then

$$A_0B_0+A_1\gamma=0 , \qquad A_0+A_1B_0'=0 .$$

If $A_1=0$, then $A_0=0$, contrary to hypothesis. Elimination of A_0 gives $A_1(\gamma-B_0'B_0)=0$.

THEOREM 6. *If $p=2$ and if Σ is a division algebra, then Γ is a division algebra if and only if**

* Note that $X'X$ and XX' are usually distinct. If they are equal when $X=j_1$, then $\alpha(\theta_1)=\alpha$. Write χ_3 for $g\gamma\alpha(\theta_1)$ and χ_2 for γ. Then $g=g(\theta_1)$ implies $\chi_3(\theta_1):\chi_3=\chi_2(\theta_1):\chi_2$. But in Cecioni's example (his §24), $\chi_3(\theta_1)=-\chi_3$, while $\chi_2(\theta_1)$ is distinct from $-\chi_2$ since $g_2\neq 0$ by his (52).

$$\gamma \neq X'X \text{ for any } X \text{ in } \Sigma \,. \tag{38}$$

If $q=2$, but p is arbitrary, Theorem 1 shows that Σ is the algebra over F_1 with the basal units $1, i, j_1, ij_1$ and is of type D with $n=2$, where F_1 is derived from F by the adjunction of the sum and the product of i and $\theta_1(i)$. It is associative if and only if $j_1^2=g$ is in F_1, i.e., if $g=g(\theta_1)$. We may apply Theorem 6 with Γ replaced by Σ and Σ replaced by the division algebra $F_1(i)$. When Σ is regarded as an algebra Γ, its q is 1, whence X' is $X(\theta_1)$. Hence for $q=2$, Σ is a division algebra if and only if $g\neq X(\theta_1)X$ for any X in $F_1(i)$, or in other words if g is not the norm, relative to F_1, of any number of $F_1(i)=F(i)$.

Applying also Theorem 5 for $p=q=2$, we obtain

THEOREM 7. *Let a quartic equation irreducible in F have the roots i, $\theta_1(i)$, $\theta_2(i)$, $\theta_3(i)$, where the $\theta_k(i)$ are rational functions of i with coefficients in F such that*

$$\theta_1\theta_1=\theta_2\theta_2=\theta_3\theta_3=i\,, \qquad \theta_1\theta_2=\theta_2\theta_1=\theta_3\,, \qquad \theta_1\theta_3=\theta_3\theta_1=\theta_2\,, \qquad \theta_2\theta_3=\theta_3\theta_2=\theta_1\,,$$

where $\theta_r\theta_s$ denotes $\theta_r[\theta_s(i)]$. Under the law of multiplication

$$(a+bj_1)(c+dj_1)=ac+gbd(\theta_1)+[ad+bc(\theta_1)]j_1\,, \tag{39}$$

where a,b,c,d,g are in $F(i)$, the elements $a+bj_1$ form an associative division algebra Σ if and only if g is in the field $F_1=F(i+\theta_1,\, i\theta_1)$, but is not the norm, relative to F_1, of any number of $F_1(i)=F(i)$. For $A=a+bj_1$, write $A'=a(\theta_2)+b(\theta_2)\alpha j_1$, where α and γ are in $F(i)$. Under the law of multiplication (37), *where A_0, A_1, B_0, B_1 are in Σ, the elements $A_0+A_1j_2$ form an associative division algebra Γ, when Σ is one, if and only if* (38) *holds and*

$$\alpha\alpha(\theta_1)g=g(\theta_2)\,, \qquad \alpha\alpha(\theta_2)\gamma(\theta_1)=\gamma=\gamma(\theta_2)\,. \tag{40}$$

This Γ is an algebra over F with the 16 basal units $i^rj_s (r,s=0,1,2,3)$, where $j_3=j_1j_2$. Special cases of our laws of multiplication give

$$\begin{gathered} j_1^2=g\,, \quad j_2^2=\gamma\,, \quad j_1j_2=j_3\,, \quad j_2j_1=\alpha j_3\,, \quad j_1j_3=gj_2\,, \\ j_2j_3=\alpha\gamma(\theta_1)j_1\,, \quad j_3j_1=g\alpha(\theta_1)j_2\,, \quad j_3j_2=\gamma(\theta_1)j_1\,, \quad j_3^2=g\gamma\alpha(\theta_1)\,, \end{gathered} \tag{41}$$

as well as $j_ri=\theta_r(i)j_r$ and its consequence (5).

Cecioni gave a special example of Γ in which are satisfied the conditions that it be an associative division algebra.

13. **Division algebras Γ with $p=3$.** We assume that Σ is a division algebra. Suppose that $\mathscr{P}\mathfrak{Q}=0$, where $\mathscr{P}$ and $\mathfrak{Q}=K+Lj_q+Mj_q^2$ are elements $\neq 0$ of Γ, K, L and M being in Σ. If $M\neq 0$, we have

$$\mathscr{R}=\mathfrak{Q}(j_q^2-Z)=U+Vj_q+Wj_q^2\,, \quad W=K-MZ''\,,$$

since $j_q^3=\gamma$ is in Σ. Then $W=0$ if $Z=(M'\gamma)^{-1}K'\gamma$. We postpone the subcase $\mathscr{R}=0$. But if $M=0$, write $\mathscr{R}=\mathfrak{Q}$. In either case, we have $\mathscr{P}\mathscr{R}=0$, where $\mathscr{R}=C+Dj_q$, $\mathscr{P}\neq 0$, $\mathscr{R}\neq 0$. If $D=0$, then $C\neq 0$, $\mathscr{P}\mathscr{R}C^{-1}=\mathscr{P}=0$. Hence $D\neq 0$, $\mathscr{A}\mathscr{B}=0$, where $\mathscr{A}=\mathscr{P}D\neq 0$, $\mathscr{B}=D^{-1}C+j_q$. Employing the notations (12) and (19) for $\mathscr{A}$ and $\mathscr{B}$, we have $B_1=1$, $B_2=0$. For $p=3$, (20) gives

$$
\begin{aligned}
&P_0=A_0B_0+A_1B_2'\gamma+A_2B_1''\gamma\ ,\\
(42)\qquad &P_1=A_0B_1+A_1B_0'+A_2B_2''\gamma\ ,\qquad P_2=A_0B_2+A_1B_1'+A_2B_0''.
\end{aligned}
$$

Since $B_1=1$, $B_2=0$, and each $P_k=0$, we have

$$A_0B_0+A_2\gamma=0\ ,\qquad A_0+A_1B_0'=0\ ,\qquad A_1+A_2B_0''=0\ .$$

Hence $A_2\neq 0$. Elimination of A_0 and A_1 gives

$$A_2(B_0''B_0'B_0+\gamma)=0\ .$$

Write $B_0=-X$. Thus $\gamma=X''X'X$. In the postponed case $\mathscr{R}=0$, we have $\mathscr{R}j_q=0$, or $\mathfrak{Q}(\gamma-Zj_q)=0$, while neither factor is zero. But this is the case $\mathscr{A}\mathscr{B}=0$ just treated.

THEOREM 8. *If $p=3$ and if Σ is a division algebra, then Γ is a division algebra if and only if*

$$\gamma\neq X''X'X \text{ for any } X \text{ in } \Sigma\ . \tag{43}$$

If $q=3$, but p is arbitrary, and if F_1 is derived from F by the adjunction of the elementary symmetric functions of i, $\theta_1(i)$, $\theta_2(i)$, Theorem 1 shows that Σ is an algebra over F_1 of type D with the basal units $i^rj_s(r,s=0,1,2)$, where $j_2=j_1^2$. It is associative if and only if $j_1^3=g$ is in F_1. We may apply Theorem 8 with Γ and Σ replaced by Σ and $F_1(i)$ respectively. When Σ is regarded as an algebra Γ, its q is 1, whence X' is $X(\theta_1)$. Hence for $q=3$, Σ is a division algebra if and only if $g\neq X(\theta_2)X(\theta_1)X$ for any X in $F_1(i)$, or in other words if g is not the norm, relative to F_1, of any number of $F_1(i)=F(i)$.

Applying also Theorem 5 for $p=q=3$, we obtain

THEOREM 9. *Let an equation of degree 9 irreducible in F have the roots*

$$\theta_1^k[\theta_3^r(i)]=\theta_3^r[\theta_1^k(i)]\qquad\qquad (k,\ r=0,\ 1,\ 2)\ ,$$

where θ_1 and θ_3 are rational functions of i with coefficients in F such that $\theta_1^3(i)=i$, $\theta_3^3(i)=i$. Let F_1 be derived from F by adjoining the elementary symmetric functions of i, θ_1, $\theta_1^2(i)$. Let Σ be the set of all linear functions of $1, j_1, j_2=j_1^2$ with

coefficients in $F(i)$ *and let multiplication be performed in* Σ *by means of* (5) *and* $j_1^3=g$. *Then* Σ *is an associative division algebra if and only if* g *is in* F_1, *but is not the norm, relative to* F_1, *of any number of* $F(i)$. *Let*

$$A'=a(\theta_3)+b(\theta_3)\alpha j_1+c(\theta_3)\alpha\alpha(\theta_1)j_2 \quad \text{for} \quad A=a+bj_1+cj_2\,. \tag{44}$$

Under the law of multiplication

$$(A_0+A_1j_3+A_2j_3^2)(B_0+B_1j_3+B_2j_3^2)=P_0+P_1j_3+P_2j_3^2\,,$$

where the P_k *are defined by* (42), *with* γ *in* $F(i)$, *the elements* $A_0+A_1j_3+A_2j_3^2$ *in which the* A_k *range independently over* Σ *form an associative division algebra* Γ *if and only if* $\gamma\neq X''X'X$ *for any* X *in* Σ *and*

$$g=g(\theta_1)\,,\quad \gamma=\gamma(\theta_3)\,,\quad \alpha\alpha(\theta_1)\alpha(\theta_1^2)g=g(\theta_3)\,,\quad \alpha\alpha(\theta_3)\alpha(\theta_3^2)\gamma(\theta_1)=\gamma\,. \tag{45}$$

Evidently Γ is an algebra over F with the 81 basal units i^rj_s $(r,s=0,1,\cdots,8)$. For brevity write f_k for $f(\theta_k)$, and $C=\alpha\alpha_1\alpha_3\alpha_4$. Then the multiplication table of Γ is given by (5) and the following relations in which j_r is denoted by r:

$$1^2=2,\quad 12=g,\quad 13=4,\quad 14=5,\quad 15=g3,\quad 16=7,\quad 17=8,\quad 18=g6,$$
$$21=g,\quad 2^2=g1,\quad 23=5,\quad 24=g3,\quad 25=g4,\quad 26=8,\quad 27=g6,\quad 28=g7,$$
$$31=\alpha 4,\quad 32=\alpha\alpha_1 5,\quad 3^2=6,\quad 34=\alpha 7,\quad 35=\alpha\alpha_1 8,\quad 36=\gamma,\quad 37=\alpha\gamma_1 1,\quad 38=\alpha\alpha_1\gamma_2 2,$$
$$41=\alpha_1 5,\quad 42=g\alpha_1\alpha_2 3,\quad 43=7,\quad 44=\alpha_1 8,\quad 45=g\alpha_1\alpha_2 6,\quad 46=\gamma_1 1,\quad 47=\alpha_1\gamma_2 2,$$
$$48=g\gamma\alpha_1\alpha_2,\quad 51=g\alpha_2 3,\quad 52=g\alpha\alpha_2 4,\quad 53=8,\quad 54=g\alpha_2 6,\quad 5^2=g\alpha\alpha_2 7,$$
$$56=\gamma_2 2,\quad 57=g\gamma\alpha_2,\quad 58=g\gamma_1\alpha\alpha_2 1,\quad 61=\alpha\alpha_3 7,\quad 62=C8,\quad 63=\gamma,$$
$$64=\alpha\alpha_3\gamma_1 1,\quad 65=C\gamma_2 2,\quad 6^2=\gamma 3,\quad 67=\alpha\alpha_3\gamma_1 4,\quad 68=C\gamma_2 5,$$
$$71=\alpha_1\alpha_4 8,\quad 72=gC_1 6,\quad 73=\gamma_1 1,\quad 74=\alpha_1\alpha_4\gamma_2 2,\quad 75=g\gamma C_1,\quad 76=\gamma_1 4,$$
$$7^2=\alpha_1\alpha_4\gamma_2 5,\quad 78=g\gamma C_1 3,\quad 81=g\alpha_2\alpha_5 6,\quad 82=gC_2 7,\quad 83=\gamma_2 2,$$
$$84=g\gamma\alpha_2\alpha_5,\quad 85=g\gamma_1 C_2 1,\quad 86=\gamma_2 5,\quad 87=g\gamma\alpha_2\alpha_5 3,\quad 8^2=g\gamma_1 C_2 4\,.$$

14. **Division algebras Γ with $p>3$.** If Γ is a division algebra, then

$$\gamma\neq X^{(p-1)}X^{(p-2)}\cdots X'X \text{ for any } X \text{ in } \Sigma\,. \tag{46}$$

For, if $\gamma=X^{(p-1)}\cdots X$, then $\mathscr{A}(j_q-X)=0$, while neither factor is zero, if

$$\mathscr{A}=\sum_{k=0}^{p-2} X^{(p-1)}X^{(p-2)}\cdots X^{(k+1)}j_q^k+j_q^{p-1}\,.$$

For $p=2$ and $p=3$, we proved that conversely (46) implies that Γ is a division algebra. But for $p>3$ no attempt is made here to prove this con-

verse. We cited in §4 Wedderburn's proof for the case in which Γ is an algebra of type D; he employed the determinant of matrix (35) in the matric representation of D. But for the corresponding matrix (21) for Γ, the notion of a determinant is absent since the elements of (21) are not commutative. This particular difficulty is overcome if we use the representation in §11 of Γ as an algebra of matrices $\{\mathscr{A}\}$ with elements in $F(i)$, whence the elements are now commutative. While the constants of multiplication of D involve a single parameter g, those of Γ involve not only the corresponding parameters g and γ, but also a parameter α connected with them by relations (28) and (29). Waiving the difficulty* which thus arises in assuming that γ is an independent variable which may be made zero without altering the analogue of Wedderburn's δ, let us attempt to apply his proof to matrix (36). Since $\gamma=\gamma(\theta_2)$, also $\delta=\delta(\theta_2)$; otherwise $\gamma+\delta$ is not zero and has an inverse. Hence in matrix $\{\gamma+\delta\}$ the elements outside the diagonal are all zero, while those in the diagonal are $e=\gamma+\delta$ and $f=\gamma(\theta_1)+\delta(\theta_1)$ each taken twice. Since $A_1=1$, we have $c=1$, $d=0$ in (36), whose determinant is the sum of $\gamma\gamma_1$ and a linear function of γ and γ_1. This determinant must divide e^2f^2 and hence is equal to ef. For $\gamma=0$, it is seen by inspection to be the product of $\rho=aa_1-gbb_1$ by $\rho(\theta_2)=a_2a_3-gb_2b_3$ in view of (40_1). Taking $\gamma=0$, we get $\delta\delta(\theta_1)=\rho\rho(\theta_2)$. But $\gamma+\delta=0$. Hence the condition is that $\gamma\gamma(\theta_1)\neq\rho\rho(\theta_2)$ for any ρ in $F(i)$.

But this is not a necessary condition that Γ be a division algebra. For, in Cecioni's example (his §24), with $\gamma=\chi_2$, we find that $\gamma\gamma(\theta_1)=-k_1^2=\rho\rho(\theta_2)$ for $\rho=k_1i$. Hence no obvious modification of the method used for D will succeed for Γ.

For $p=2$, q arbitrary, the corresponding condition is

$$\gamma\gamma(\theta_1)\cdot\cdot\cdot\gamma(\theta_{q-1})\neq\sigma\sigma(\theta_q) \text{ for any } \sigma \text{ in } F(i)\,.$$

15. **Algebras Γ whose group G is not necessarily abelian.** As at the end of §5, let G have an invariant subgroup G_q composed of $\Theta_0=1, \Theta_1, \cdots, \Theta_{q-1}$ and let G_q be extended to G by Θ_q. If p is the index of G_q under G, then Θ_q^p is a substitution Θ_e of G_q, while no lower than the pth power of Θ_q belongs to G_q. Also,

$$\Theta_k\Theta_q=\Theta_q\Theta_{k_0} \qquad (k<q)\,, \tag{47}$$

and the substitutions of G are given without repetition by

$$\Theta_{rq+k}=\Theta_q^r\Theta_k \qquad (r=0,1,\cdots,p-1\,;\ k=0,1,\cdots,q-1)\,. \tag{48}$$

* We assume that $q=2$. In $(A+j_2)(B+j_2)=\gamma+AB+(A+B')j_2$, we have $A+B'=0$ by hypothesis. Then $\delta=AB=-B'B$ involves α and hence depends on γ.

Hence (9) holds. As in §7, we may take

$$(49)\quad j_q^r=j_{rq}\ ,\quad j_kj_{rq}=j_{k+rq}\qquad (r=1,\cdots,p-1\ ;\ k=1,\cdots,q-1)$$

and conclude that every element of Γ is of the form (12). By (47) and (6),

$$(50)\quad \theta_k[\theta_q(i)]=\theta_q[\theta_{k_0}(i)]\ ,\quad j_qj_k=\alpha_kj_{k_0}j_q\qquad (k=1,\cdots,q-1)\ ,$$

where α_k is in $F(i)$. Thus

$$(51)\quad j_qA=A'j_q\ ,\quad A'=f_0(\theta_q)+\sum_{k=1}^{q-1}f_k(\theta_q)\alpha_kj_{k_0}\quad \text{for } A \text{ in (13)}\ .$$

Since $\Theta_q^p=\Theta_e$,

$$(52)\quad \theta_q^p(i)=\theta_e(i)\ ,\qquad j_q^p=\gamma\equiv\beta j_e\ ,$$

where β is a number $\neq 0$ of $F(i)$, and $e<q$. Since j_q is commutative with j_q^p,

$$\beta j_{e+q}=\beta j_ej_q=j_q\beta j_e=\beta(\theta_q)j_qj_e=\beta(\theta_q)\alpha_ej_{e_0}j_q\ ,$$

and the final product of the j's is j_{e_0+q}. Hence $e_0=e$ and

$$(53)\quad \beta=\beta(\theta_q)\alpha_e\ ,$$

where $\alpha_e=1$ if $e=0$. Hence $\gamma'=\gamma$. We have (18) and

$$(54)\quad A^{(p)}\gamma=\gamma A\ .$$

By Theorem 2, these relations imply that Γ is associative. We now investigate the conditions for these relations. We require that $(AB)'=A'B'$ for $A=fj_k$, $B=hj_r$ for all f and h in $F(i)$ and for $k, r=0, 1, \cdots, q-1$. We may write $\theta_r[\theta_k(i)]=\theta_u(i)$, whence $j_kj_r=c_{kr}j_u$. Replace i by $\theta_q(i)$ and apply (50_1) three times; we get

$$\theta_r\theta_k\theta_q=\theta_r\theta_q\theta_{k_0}=\theta_q\theta_{r_0}\theta_{k_0}=\theta_u\theta_q=\theta_q\theta_{u_0}\ ,$$

whence

$$\theta_{r_0}\theta_{k_0}=\theta_{u_0}\ ,\qquad j_{k_0}j_{r_0}=c_{k_0r_0}j_{u_0}\ .$$

Then

$$A'=f(\theta_q)\alpha_kj_{k_0}\ ,\qquad B'=h(\theta_q)\alpha_rj_{r_0}\ ,\qquad AB=fh(\theta_k)c_{kr}j_u\ ,$$

$$A'B'=f(\theta_q)\alpha_kh[\theta_q\{\theta_{k_0}(i)\}]\alpha_r(\theta_{k_0})c_{k_0r_0}j_{u_0}\ ,$$

$$(AB)'=f(\theta_q)h[\theta_k\{\theta_q(i)\}]c_{kr}(\theta_q)\alpha_uj_{u_0}\ .$$

The two h's are equal by (50). Hence the conditions are

$$(55)\quad \alpha_k\alpha_r(\theta_{k_0})c_{k_0r_0}=c_{kr}(\theta_q)\alpha_u\qquad (k,\ r=1,\cdots,q-1\ ;\ \alpha_0=1)\ ,$$

being satisfied identically if $k=0$ or $r=0$.

We next seek the conditions for (54). Write k_{00} for $(k_0)_0$, etc. By (47) we find by induction that

$$\Theta_k \Theta_q^s = \Theta_q^s \Theta_{k_0 \cdots 0}, \tag{56}$$

where the number of zeros is s. The case $s=p$ gives

$$\theta_k \theta_e = \theta_e \theta_{k_0 \cdots 0}, \qquad j_e j_k = d_k j_{k_0 \cdots 0} j_e, \tag{57}$$

where there are p subscripts 0 to k, and d_k is in $F(i)$. By induction on s,

$$A^{(s)} = f_0(\theta_q^s) + \sum_{k=1}^{q-1} f_k(\theta_q^s)\alpha_k(\theta_q^{s-1})\alpha_{k_0}(\theta_q^{s-2})\alpha_{k_{00}}(\theta_q^{s-3}) \cdots \alpha_{k_0 \cdots 0} j_{k_0 \cdots 0},$$

where there are $s-1$ subscripts 0 to k under the final α, and s of them under j. Take $s=p$ and apply (52_1). We see that the desired conditions for (54) are

$$\beta d_k = \alpha_k(\theta_q^{p-1})\alpha_{k_0}(\theta_q^{p-2})\alpha_{k_{00}}(\theta_q^{p-3}) \cdots \alpha_{k_0 \cdots 0}\beta(\theta_{k_0 \cdots 0}) \qquad (k=1, \cdots, q-1), \tag{58}$$

where there are $p-1$ subscripts 0 under the last α, and p under the final θ.

THEOREM 10. *If the subalgebra with the units $i^r j_k$ $(r<pq, k<q)$ is associative, then Γ is associative if and only if conditions* (53), (55), (58) *hold.*

Consider the simplest non-abelian case in which G_q is a cyclic group generated by Θ_1 such that Θ_q transforms Θ_1 into its inverse. Since Θ_q transforms $\Theta_k = \Theta_1^k$ into Θ_{q-k} and into Θ_{k_0} by (47), the latter subscripts differ only by a multiple of q. Hence

$$\text{if } k=0, \quad k_0=0; \quad \text{if } k>0, \quad k_0=q-k. \tag{59}$$

We employ the values of c_{rk} and u given above ((25) in §10). In (55), $k_0=q-k$, $r_0=q-r$. If $r+k<q$, then $r_0+k_0>q$ and (55) becomes

$$\alpha_k \alpha_r(\theta_1^{q-k}) g = \alpha_{k+r} \qquad (r, k=1, \cdots, q-1; \ r+k<q). \tag{60}$$

For $r=1$ this gives by induction on k

$$\alpha_k = \alpha\alpha(\theta_1^{q-1})\alpha(\theta_1^{q-2}) \cdots \alpha(\theta_1^{q-k+1}) g^{k-1} \qquad (k=1, \cdots, q-1), \tag{61}$$

where $\alpha=\alpha_1$. Then (60) is seen to be satisfied when we insert the values of α_k, α_r, α_{k+r} from (61) and apply $g(\theta_1)=g$, $\theta_1^{q+s}=\theta_1^s$.

For $r+k=q$, whence $r_0+k_0=q$, (55) becomes

$$\alpha_k \alpha_r(\theta_1^{q-k}) g = g(\theta_q). \tag{62}$$

Inserting the values of α_k and α_r from (61), we get

$$\alpha\alpha(\theta_1)\alpha(\theta_1^2)\cdots\alpha(\theta_1^{q-1})g^{q-1}=g(\theta_q)\,. \tag{63}$$

Finally, for $r+k>q$, (55) becomes

$$\alpha_k\alpha_r(\theta_1^{q-k})=g(\theta_q)\alpha_{k+r-q}\,. \tag{64}$$

Inserting the values (61) of the three α's, we see that the α's from the third α all cancel a like number of α's in the new left member, and that the resulting relation is (63).

Since Θ_q transforms Θ_1 into its inverse and vice versa, Θ_q^s transforms Θ_1 into itself or its inverse according as s is even or odd. Take $s=p$ and note that $\Theta_q^p=\Theta_e=\Theta_1^e$ transforms Θ_1 into itself. We exclude the case $q=2$ since Θ_q then transforms Θ_1 into itself and G is abelian. Hence p is even.

By (59), $k_{00}=k$. Thus (57) becomes $j_e j_k=d_k j_k j_e$. But $j_e=j_1^e, j_k=j_1^k$. Hence $d_k=1$. Thus (58) is

$$\beta=\alpha_k(\theta_q^{p-1})\alpha_{q-k}(\theta_q^{p-2})\alpha_k(\theta_q^{p-3})\alpha_{q-k}(\theta_q^{p-4})\cdots\alpha_k(\theta_q)\alpha_{q-k}\beta(\theta_k)\,, \tag{65}$$

for $k=1,\cdots,q-1$. We need the formula

$$\alpha_{q-1}\alpha_{q-1}(\theta_1)\cdots\alpha_{q-1}(\theta_1^{t-1})=[g(\theta_q)]^{t-1}\alpha_{q-t}\,. \tag{66}$$

To prove it by induction on t, multiply it by $\alpha_{q-1}(\theta_1^t)$ and apply (64) for $k=q-t$, $r=q-1$. We shall prove that (65) follows from the case $k=1$ by replacing i by $i, \theta_1, \cdots, \theta_1^{k-1}$ in turn and taking the product. It suffices to watch the product of the general pair of consecutive factors

$$\alpha_1(\theta_q^{p-2s+1})\alpha_{q-1}(\theta_q^{p-2s})\,.$$

Replace i by θ_1^c and apply

$$\theta_q^{p-2s+1}\theta_1^c=\theta_1^{q-c}\theta_q^{p-2s+1}\,,\qquad \theta_q^{p-2s}\theta_1^c=\theta_1^c\theta_q^{p-2s}\,,$$

which follows from (56) which states that $\theta_1^k\theta_q^s=\theta_q^s\theta_1^k$ or $\theta_q^s\theta_1^{q-k}$, according as s is even or odd. Hence we get

$$\prod_{c=0}^{k-1}\alpha_1(\theta_1^{q-c}\theta_q^{p-2s+1})\cdot\prod_{c=0}^{k-1}\alpha_{q-1}(\theta_1^c\theta_q^{p-2s})\,.$$

By (61), the first product is derived from α_k/g^{k-1} by replacing i by θ_q^{p-2s+1}. By (66) with $t=k$, the second product is derived from $[g(\theta_q)]^{k-1}\alpha_{q-k}$ by replacing i by θ_q^{p-2s}. Thus the g's cancel and we get

$$\alpha_k(\theta_q^{p-2s+1})\alpha_{q-k}(\theta_q^{p-2s})\,,$$

which is the product of the corresponding pair of consecutive factors in (65). Hence all cases of (65) follow from the case $k=1$. We first state our results for the case $e=0$, whence $\beta=\gamma$.

THEOREM 11. *Let $f(x)=0$ be an equation of degree pq irreducible in F whose Galois group G for F is generated by two substitutions Θ_1 and Θ_q of respective orders q and p such that Θ_q transforms Θ_1 into its inverse, while no lower than the pth power of Θ_q is equal to a power of Θ_1. Excluding the case $q=2$, we see that G is not abelian and that p is even. Then the roots of $f(x)=0$ are*

$$\theta_q^r[\theta_1^k(i)]=\begin{matrix}\theta_1^k[\theta_q^r(i)] & (r\ even) \\ \theta_1^{q-k}[\theta_q^r(i)] & (r\ odd)\end{matrix}\qquad\begin{pmatrix} r=0,1,\cdots,p-1 \\ k=0,1,\cdots,q-1\end{pmatrix},\tag{67}$$

where θ_1 and θ_q are rational functions of i with coefficients in F such that the qth iterative $\theta_1^q(i)$ of $\theta_1(i)$ is i and likewise $\theta_q^p(i)=i$. There exists an associative algebra Σ whose elements are

$$A=f_0+f_1j_1+f_2j_1^2+\cdots+f_{q-1}j_1^{q-1}\ ,\tag{68}$$

where the f_k are polynomials in i of degree $<pq$ with coefficients in F, while

$$j_1^q=g(i)=g(\theta_1)\ ,\qquad j_1^r\phi(i)=\phi[\theta_1^r(i)]j_1^r\qquad(r=1,\cdots,q-1)\ ,\tag{69}$$

so that the product of any two elements (68) *of Σ is another element* (68) *of Σ. Let*

$$A'=f_0(\theta_q)+\sum_{k=1}^{q-1}f_k(\theta_q)\alpha_k j_{q-k}\ ,$$

where α_k is defined by (61). *Then under multiplication defined by* (20), *the totality of polynomials in j_q with coefficients in Σ form an algebra Γ of order p^2q^2 over F which is associative if and only if $\gamma=\gamma(\theta_q)$,*

$$\gamma=\alpha(\theta_q^{p-1})\alpha_{q-1}(\theta_q^{p-2})\alpha(\theta_q^{p-3})\alpha_{q-1}(\theta_q^{p-4})\cdots\alpha(\theta_q)\alpha_{q-1}\gamma(\theta_1)\ ,\tag{70}$$

and also (63) *holds. Hence there are only four conditions on the parameters g, γ, α of Γ.*

The associative conditions are not inconsistent, being all satisfied if γ is F, and $g(\theta_1)=g$, $g(\theta_q)=g^{-1}$, $\alpha=g^{-1}$, whence every $\alpha_k(\theta_q^s)$ is g or g^{-1} according as s is odd or even.

For example, let $p=2$, $q=3$, and let

$$\Theta_1=(0\ 1\ 2)(3\ 5\ 4)\ ,\qquad \Theta_3=(0\ 3)(1\ 4)(2\ 5)\ .$$

Then the roots of the sextic are

$$i,\ \theta_1,\ \theta_2=\theta_1[\theta_1(i)]\ ,\quad \theta_3,\ \theta_4=\theta_3[\theta_1(i)]=\theta_2[\theta_3(i)]\ ,\quad \theta_5=\theta_3\theta_2=\theta_1\theta_3\ .$$

For brevity write f_k for $f(\theta_k)$ and $\beta=\alpha\alpha(\theta_2)g$ for our former α_2. Then the multiplication table of Γ is given by (69_2) and the following relations in which j_r is denoted by r:

$$
\begin{array}{lllll}
1^2=2\ , & 12=g\ , & 13=4\ , & 14=5\ , & 15=g3\ , \\
21=g\ , & 2^2=g1\ , & 23=5\ , & 24=g3\ , & 25=g4\ , \\
31=\alpha 5\ , & 32=\beta 4\ , & 3^2=\gamma\ , & 34=\alpha\gamma_2 2\ , & 35=\beta\gamma_1 1\ , \\
41=\alpha_1 g3\ , & 42=\beta_1 5\ , & 43=\gamma_1 1\ , & 4^2=\alpha_1\gamma g\ , & 45=\beta_1\gamma_2 2\ , \\
51=\alpha_2 g4\ , & 52=\beta_2 g3\ , & 53=\gamma_2 2\ , & 54=\alpha_2\gamma_1 g1\ , & 5^2=\beta_2\gamma g\ .
\end{array}
$$

The conditions for associativity are

$$g_1=g\ , \quad \gamma_3=\gamma\ , \quad g_3=\alpha\alpha_1\alpha_2 g^2\ , \quad \gamma=\alpha\alpha_2\alpha_3 g\gamma_1\ .$$

Let F_1 be the field obtained from F by adjoining the elementary symmetric functions of $1, \theta_1, \theta_2$. Then Γ is a division algebra if and only if g is not the norm, relative to F_1, of any number of $F(i)$, and if $\gamma\neq X'X$ for any $X=a+bj_1+cj_1^2$, a, b, c in $F(i)$. Here

$$X'=a_3+c_3\alpha\alpha_2 g j_1+b_3\alpha j_2\ .$$

For any p and q, p even, we readily exhibit the group G of Theorem 11 in the desired regular form. Take $\Theta_1=C_1C_2\cdots C_p$, where each C_i is a cycle of q letters, no two C's having a letter in common. Write $\Theta_1^{-1}=C_2^{-1}C_3^{-1}\cdots C_p^{-1}C_1^{-1}$, where the cycle C_i^{-1} starts with the same letter as C_i, so that its remaining letters are those of C_i taken in reverse order. Then Θ_q is the substitution which replaces each letter of Θ_1 by that in the corresponding position in Θ_1^{-1}.

For example, if $p=2$, take $C_1=(0\ 1\ \cdots\ q-1)$, $C_2=(q\ q+1\ \cdots\ 2q-1)$. Then

$$C_2^{-1}=(q\ \ 2q-1\ \ 2q-2\ \cdots\ q+2\ \ q+1)\ , \qquad C_1^{-1}=(0\ \ q-1\ \ q-2\ \cdots\ 2\ \ 1)\ ,$$

$$\Theta_q=(0\ \ q)(1\ \ 2q-1)(2\ \ 2q-2)\cdots(q-1\ \ q+1)\ .$$

For $p=4$, $q=3$,

$$\Theta_1=(0\ 1\ 2)(3\ 4\ 5)(6\ 7\ 8)(9\ 10\ 11)\ ,$$

$$\Theta_1^{-1}=(3\ 5\ 4)(6\ 8\ 7)(9\ 11\ 10)(0\ 2\ 1)\ ,$$

$$\Theta_3=(0\ 3\ 6\ 9)(1\ 5\ 7\ 11)(2\ 4\ 8\ 10)\ .$$

Next, let $e>0$. Since $e_0=e$, we have $2e=q$ by (59). Thus p and q are now both even.

THEOREM 12. *Let $f(x)=0$ be an equation of degree pq irreducible in F whose Galois group G is generated by Θ_1 and Θ_q, such that Θ_1 is of order q, Θ_q*

transforms Θ_1 *into its inverse,* $\Theta_q^p=\Theta_1^{q/2}$, *while no lower than the pth power of* Θ_q *is equal to a power of* Θ_1. *Excluding the case* $q=2$, *we see that G is not abelian and that p and q are both even. The roots are given by* (67), *where* $\theta_1^q(i)=i$, $\theta_q^p(i)=\theta_1^{q/2}(i)$. *Consider the algebras* Σ *and* Γ *of Theorem* 11, *where now* $j_q^p=\gamma=\beta j_{q/2}$. *Then* Γ *is associative if and only if* $g=g(\theta_1), \beta=\beta(\theta_q)\alpha_{q/2}$, (63) *holds, and* (65) *holds when* $k=1$.

The associative conditions are not inconsistent, being all satisfied if β is in F, and $g(\theta_1)=g$, $g(\theta_q)=g^{-1}$, $\alpha=g^{-1}$.

The simplest example is given by $p=2$, $q=4$. We may take*

$$\Theta_1=(0\ 1\ 2\ 3)(4\ 5\ 6\ 7)\ , \qquad \Theta_4=(0\ 4\ 2\ 6)(1\ 7\ 3\ 5)\ .$$

For any even integers p and q, we readily exhibit the group of Theorem 12 in the desired regular form. The only modification to make in the method used under Theorem 11 is that we start the final cycle C_1^{-1} of Θ_1^{-1} with the first letter after the middle of C_1.

For example, if $p=q=4$, take $\Theta_1=C_1C_2C_3C_4$, $C_1=(0\ 1\ 2\ 3)$, $C_2=(4\ 5\ 6\ 7)$,

$$C_3=(8\ 9\ 10\ 11),\ C_4=(12\ 13\ 14\ 15),\ \Theta_1^{-1}=C_2^{-1}C_3^{-1}C_4^{-1}C_1^{-1},$$

$$\Theta_1^{-1}=(4\ 7\ 6\ 5)(8\ 11\ 10\ 9)(12\ 15\ 14\ 13)(2\ 1\ 0\ 3)\ .$$

The substitution which replaces each letter of Θ_1 by that in the corresponding position in Θ_1^{-1} is

$$\Theta_4=(0\ 4\ 8\ 12\ 2\ 6\ 10\ 14)(1\ 7\ 9\ 15\ 3\ 5\ 11\ 13)\ .$$

Then $\Theta_4^4=\Theta_1^2$. The groups with $p=2$ are called dicyclic.

---

* This group is the regular form of the group whose 8 elements are the quaternions ± 1, $\pm i$, $\pm j$, $\pm k$.

THE UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

THE FIRST AND SECOND VARIATIONS OF A DOUBLE INTEGRAL FOR THE CASE OF VARIABLE LIMITS*

BY

H. A. SIMMONS

Introduction

Among the many interesting general problems in the calculus of variations is the following one. Suppose we have given the double integral

$$I=\iint_{A_0} f(x,y,z,p,q)\,dx\,dy,$$

where A_0 is the area in the xy-plane over which the integral is taken (see Figure 1), and where $p\equiv\partial z/\partial x$, $q\equiv\partial z/\partial y$, and suppose we also have given a fixed surface $\varphi(x,y,z)=0$ which is arbitrary. It is then required to find among all surfaces which are representable in the form $z=z(x,y)$ and which have their edges on the fixed surface $\varphi=0$ that one which minimizes the double integral I.

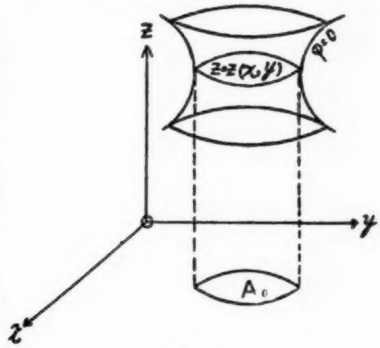

Fig. 1

In studying the problem just stated, one immediately finds need of the first and second variations of the integral I. In obtaining them, it is desirable to assume that the sought surface $z=z(x,y)$ has already been found and then to consider a one-parameter family of surfaces of the type

$$z=z(x,y)+a\zeta(x,y)$$

containing the surface $z=z(x,y)$ for $a=0$. We allow ζ to be an arbitrary function of x and y, but we take the value of the integral only on the portion of each surface of the family which is bounded by the fixed surface $\varphi=0$ (see Figure 1). When z in the integrand of the double integral I is replaced

---

* Presented to the Society, April 2, 1926; received by the editors before October, 1925.

by $z+a\zeta$, I becomes a function of a and of course the limits of the integral also change. We may then write in place of the integral above

$$I(a)=\iint_{A_a} f(x,y,z+a\zeta,p+a\zeta_x,q+a\zeta_y)\,dx\,dy,$$

where A_a is the area bounded by the projection on the xy-plane of the curve of intersection of the surface $\varphi=0$ with that surface of the family $z=z(x,y)+a\zeta(x,y)$ whose parameter value is a. Proceeding in this manner and then differentiating $I(a)$ with respect to a and putting $a=0$, we obtain the two derivatives $I'(0)$, $I''(0)$, which are of great service in studying the problem.

Heretofore, only the first derivative,* $I'(0)$, has been discussed in detail for problems of this type for which the boundary of the surfaces $z=z(x,y)$ over which the integral I is taken is allowed to vary arbitrarily on a fixed surface $\varphi=0$. The expression for $I'(0)$ presented in this paper seems to be the most satisfactory one that has been obtained, the results of other writers, Sarrus for example, being unsymmetric and very difficult to apply. No one, so far as is known to the author, has computed the second derivative $I''(0)$ heretofore, apparently because of the very great complications which have arisen when the computation has been attempted. These difficulties have been described by Bolza† and mention of them has also been made in the French Encyclopédie.‡

It is the purpose of this paper to compute symmetric, usable forms for $I'(0)$ and $I''(0)$, especially $I''(0)$. The author believes that those who study the method used here, which was suggested by Professor G. A. Bliss, will agree that it is a much better mode of attack than has been used heretofore. Furthermore, the reader will without doubt appreciate the relative simplicity and the perfect symmetry of the results if he compares them with those obtainable, let us say, by the method of Sarrus.

In §1 a theorem is proved concerning the differentiation of a double integral with respect to a parameter which occurs both in the integrand and in the limits of integration. In §2 the problem of the following sections is stated in a more precise form than is given above. In §3 the first derivative $I'(0)$ is computed and from the new expression for it some results are deduced which have already been obtained by other methods. In §4, we compute $I''(0)$. It is found to be equal to a double integral with the same integrand as that which appears for the corresponding problem for which

* See Sarrus's paper in Mémoires, Savants Etrangers, vol. 10 (1866) (Prize article).

† *Vorlesungen über Variationsrechnung*, p. 669.

‡ Encyclopédie des Sciences Mathématiques, tome II, vol. 6, p. 166.

the limits are fixed, plus a line integral which involves ζ^2 but no other power or derivative of ζ. In §5, we state a new necessary condition for a minimum of the integral I. The formulation of the statement is in terms of a boundary value problem associated with the second variation $I''(0)$. In §6, finally, applications of two of our formulas are made to the case where $z=z(x, y)$ is a surface of minimum area.

The author is pleased to acknowledge the assistance of Professor G. A. Bliss, who suggested the problem studied in this paper. It was through his careful direction and inspiration, for which the author is very thankful, that this paper was written.

1. **The derivatives of a double integral with respect to a parameter.** Let C_0 be a simply closed curve with equations

$$x=\xi(u), \qquad y=\eta(u),$$

where ξ, η are defined for all real values of u, are of class C'',* have $\xi'^2+\eta'^2=1$, and have a period l equal to the length of C_0.

Introduce near C_0 a uv-coördinate system determined by the equations

$$x=\xi(u)+v\eta'(u), \qquad y=\eta(u)-v\xi'(u). \tag{1}$$

These equations define a unique pair of coördinates (u, v) for each point (x, y) near C_0, since for each pair of values (x, y) on C_0 they define a unique pair $(u, v)=(u, 0)$ $(0\leqq u\leqq l)$ and since along C_0 the functional determinant

$$\begin{vmatrix} x_u & y_u \\ x_v & y_v \end{vmatrix} = \begin{vmatrix} \xi'+v\eta'' & \eta'-v\xi'' \\ \eta' & -\xi' \end{vmatrix} = -(1+v/\rho) \tag{2}$$

is different from zero.† Literal subscripts to functions, here as in all of this paper, indicate partial differentiation with respect to the letter written; and in the last formula ρ is the radius of curvature of C_0. The u-curves are the parallel curves to C_0 and the v-curves are the normals to C_0, thus establishing near C_0 a curvilinear coördinate system u, v with orientation opposite to that of x, y.

Now consider a family of curves, one of which, C_a (see Figure 2) is given by the equations

$$x=\xi(u)+v(u, a)\eta'(u), \qquad y=\eta(u)-v(u, a)\xi'(u), \tag{3}$$

---

* A function $f(x)$ is of class $C^{(n)}$on the interval $x_0\leqq x\leqq x_1$ if its nth derivative is continuous at every point of that interval.

† For details as to the neighborhood in which the correspondence is unique, see Bliss's *Princeton Colloquium Lectures*, 1913, p. 20.

where $v(u, a)$ is defined and of class C'' for all (u, a) having u real and a sufficiently near zero; where $v(u, a)$ has for the variable u a period $L(a)$ reducing to $L(0)=l$ when $a=0$; and where $v(u, 0)\equiv 0$. All of the curves C_a are closed on account of this periodic property, and each C_a is further simply closed for all values of a sufficiently near zero since C_0 is simply closed. We let A_a denote the area bounded by C_a.

Let C_1 be the curve defined by (1) when the variable v is replaced by a special constant value $v_1<0$, and designate by A_1 the area bounded by C_1 (see Figure 2 above). In the paragraphs below, v_1 is supposed to be as near zero as may be required in order that the curve C_1 defined by the constant value v_1 shall lie in the neighborhood of C_0 in which the uv-coördinate system described above is well-defined.

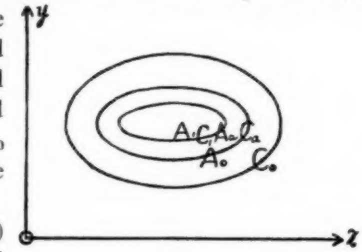

FIG. 2

Let $g(x, y, a)$ be a function of x, y, a which is of class C'' for all sets (x, y, a) having (x, y) in a neighborhood of the area A_0 bounded by C_0, and having a sufficiently near $a=0$. Let us define $J(a)$ by the formula

$$J(a)\equiv \iint_{A_a} g(x, y, a)\, dx\, dy\,. \tag{4}$$

We wish to find the derivatives $J'(0)$ and $J''(0)$ of this integral. It is convenient in order to make these differentiations to have $J(a)$, (4), written in the form

$$J(a)= \iint_{A_1} g(x, y, a)\, dx\, dy+ \iint_{\Delta A} g(x, y, a)\, dx\, dy\,, \tag{4'}$$

where A_1 is as defined above and ΔA is the area bounded by the curves C_1 and C_a (see Figure 2). The derivative of the first integral of (4′) has, as is well known, the value

$$\iint_{A_1} g_a(x, y, a)\, dx\, dy\,. \tag{5}$$

To find the derivative of the second integral in (4′), we first transform it to the uv-coördinate system by means of (1), remembering the value (2) of the functional determinant $x_u y_v - x_v y_u$, and obtain

$$\iint_{\Delta A} g(x,y,a)\, dx\, dy = \int_0^l \left[\int_{v_1}^{v(u,a)} g(\xi+v\eta', \eta-v\xi', a)(1+v/\rho)dv\right] du\,. \tag{6}$$

The derivative of the last written integral is obtainable by the usual formula for the derivative of a simple integral containing a parameter. Since the parameter a occurs only in the upper limit of the inner integral and explicitly in g, we find for this derivative

$$\int_0^l v_a g\,(1+v/\rho)\,du + \int_0^l \int_{v_1}^{v(u,a)} g_a(1+v/\rho)\,dv\,du\,,$$

where in the first integral the v, wherever it occurs, is the function $v(u, a)$. Adding this result to the expression (5), we obtain

$$J'(a) = \iint_{A_1} g_a dx\,dy + \int_0^l \int_{v_1}^{v(u,a)} g_a(1+v/\rho)\,dv\,du$$
$$+ \int_0^l v_a g\,(\xi'+v\eta', \eta-v\xi', a)\,(1+v/\rho)\,du$$

or, after transforming the second integral to xy-coördinates again,

$$J'(a) = \iint_{A_a} g_a\,dx\,dy + \int_0^l v_a\,g\,(\xi+v\eta', \eta-v\xi', a)\,(1+v/\rho)\,du\,. \tag{7}$$

To calculate the value of the second derivative $J''(a)$, we can in like manner find the derivative of the first term of (7) to be

$$\iint_{A_a} g_{aa} dx\,dy + \int_0^l g_a v_a\,(1+v/\rho)\,du\,.$$

By the formula for differentiating a simple integral with respect to a parameter, the derivative of the last term of (7) is

$$\int_0^l \left[(gv_{aa}+g_v v_a^2+g_a v_a)(1+v/\rho)+gv_a^2/\rho\right]du\,.$$

Therefore

$$J''(a) = \iint_{A_a} g_{aa}\,dx\,dy + \int_0^l \left[(gv_{aa}+2g_a v_a+g_v v_a^2)(1+v/\rho)+gv_a^2/\rho\right]du\,. \tag{8}$$

Putting $a=0$ in (7) and (8), we obtain the desired results which are described in the following theorem.

THEOREM 1. *The derivatives $J'(0)$ and $J''(0)$ of the function $J(a)$, defined by the double integral* (4), *taken over the region A_a bounded by the curve C_a, defined by the equations* (3), *have the values*

$$\begin{aligned} J'(0) &= \iint_{A_0} g_a\,dx\,dy + \int_0^l gv_a\,du\,, \\ J''(0) &= \iint_{A_0} g_{aa} dx\,dy + \int_0^l (gv_{aa}+gv_a^2/\rho+2g_a v_a+g_v v_a^2/\rho)\,du\,. \end{aligned} \tag{9}$$

Although these derivatives (9) have been computed for a one-parameter family of curves of the special type (3), we can obtain from them analogous formulas for a more general family of the form

$$x=X(\tau, a), \quad y=Y(\tau, a). \tag{10}$$

We suppose that (10) represents a one-parameter family of simply closed curves containing C_0 for $a=0$ and having τ as length of arc on C_0. The functions X, Y are of class C'' for all values of (τ, a) having τ real and the values of a sufficiently near zero. They have a period $T(a)$ for every a with $T(0)=l$, the length of C_0. Such a family is always representable in the form (3) above by solving the equations

$$\xi(u)+v\eta'(u)-X(\tau, a)=0, \quad \eta(u)-v\xi'(u)-Y(\tau, a)=0 \tag{11}$$

for v and τ as functions of u and a. According to the implicit function theorem referred to above, this can always be done since equations (11) have the particular set of solutions $(v, \tau, u, a)=(0, u, u, 0)$ for $0 \leqq u \leqq l$, on which

$$\begin{vmatrix} \eta' & -X_\tau \\ -\xi' & -Y_\tau \end{vmatrix} = \begin{vmatrix} \eta' & -\xi' \\ -\xi' & -\eta' \end{vmatrix} = -1.$$

Hence, according to the implicit function theorem referred to above, if $J(a)$ is the value of J on the region A_a bounded by the curve C_a, defined by (10), its derivatives $J'(0)$ and $J''(0)$ are given by (9), where v_a is obtained by differentiating equations (11) with respect to a and solving the resulting equations for v_a, τ_a; and where v_{aa} is similarly obtained after a second differentiation and solution of the two resulting equations, which are linear in v_{aa} and τ_{aa}, for v_{aa}. The derivatives are

$$\begin{aligned} &X_\tau\tau_a-\eta' v_a+X_a=0\,; \qquad Y_\tau\tau_a+\xi' v_a+Y_a=0\,; \\ &X_\tau\tau_{aa}-\eta' v_{aa}+X_{\tau\tau}\tau_a^2+2X_{a\tau}\tau_a+X_{aa}=0\,; \\ &Y_\tau\tau_{aa}+\xi' v_{aa}+Y_{\tau\tau}\tau_a^2+2Y_{a\tau}\tau_a+Y_{aa}=0\,. \end{aligned} \tag{12}$$

From the two equations in the first line of (12) we obtain v_a and τ_a. Then using the value found for τ_a, namely

$$\tau_a=\frac{-\xi' X_a-\eta' Y_a}{\xi' X_\tau+\eta' Y_\tau}=-X_\tau X_a-Y_\tau Y_a\,,$$

in the last two equations, we solve them, as stated above, for v_{aa}. This gives for v_a and v_{aa} the values

$$\begin{aligned} v_a&=X_aY_\tau-X_\tau Y_a\,, \\ v_{aa}&=(X_{\tau\tau}Y_\tau-X_\tau Y_{\tau\tau})(X_\tau X_a+Y_\tau Y_a)^2 \\ &\quad -2(X_{a\tau}Y_\tau-X_\tau Y_{a\tau})(X_\tau X_a+Y_\tau Y_a)+X_{aa}Y_\tau-X_\tau Y_{aa}. \end{aligned} \tag{13}$$

Substituting the values (13) of v_a and v_{aa} in equations (9), we obtain the more general result desired, which we express in the

COROLLARY. *The derivatives $J'(0)$ and $J''(0)$ of the double integral* (4), *taken over the region A_a, bounded by the curve C_a defined by the equations* (10), *have the values*

$$(14)\qquad \begin{aligned} J'(0) &= \iint_{A_0} g_a\,dx\,dy + \int_0^l g(X_aY_\tau - X_\tau Y_a)\,du\,, \\ J''(0) &= \iint_{A_0} g_{aa}\,dx\,dy + \int_0^l Q\,du, \end{aligned}$$

where

$$\begin{aligned} Q = g\big[&(X_{\tau\tau}Y_\tau - X_\tau Y_{\tau\tau})(X_\tau X_a + Y_\tau Y_a) - 2(X_{a\tau}Y_\tau - X_\tau Y_{a\tau})(X_\tau X_a + Y_\tau Y_a) \\ &+ X_{aa}Y_\tau - X_\tau Y_{aa}\big] + 2(X_aY_\tau - X_\tau Y_a)g_a + (X_aY_\tau - X_\tau Y_a)^2(g_\nu + g/\rho)\,. \end{aligned}$$

2. **Statement of the problem.** Suppose we have given an arbitrary, fixed surface

$$(15)\qquad \varphi(x, y, z) = 0\,,$$

which is of class C'' and has no singular points for all x, y, z that we wish to consider. Let the surface

$$(16)\qquad z = z(x, y)$$

be of class C'' for all x, y, z under consideration, and consider that portion of it which is bounded by its intersection with the surface (15) as indicated in Figure 3. Designate by C_0' the common line of these surfaces and by C_0 its projection on the xy-plane. Consider now the one-parameter family of surfaces

$$(17)\qquad z = z(x, y) + a\zeta(x, y)\,,$$

where ζ is an arbitrary function of x, y which is of class C'' for all x, y sufficiently near those inside or on C_0, and where a is a parameter, and where the portions of all surfaces of the family (17) to be considered are those which are bounded by the intersections of these surfaces with the fixed surface (15). These intersections determine a one-parameter family of curves which include C_0' for $a = 0$. The curve of this family with parameter a we call C_a'; its equations are

$$(18)\qquad \varphi(x, y, z) = 0, \quad z = z(x, y) + a\zeta(x, y).$$

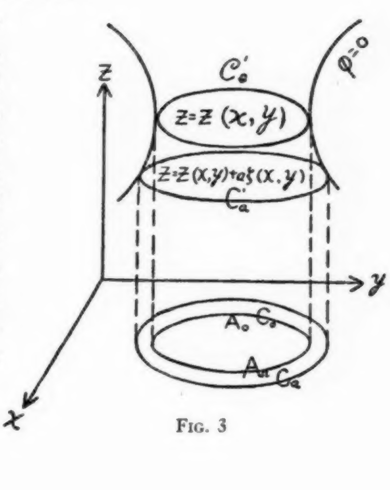

FIG. 3

Denote by C_a the projection of the xy-plane of C_a' and designate by A_a the area in the xy-plane bounded by C_a.

Consider now the double integral

$$I = \iint_{A_0} f(x, y, z, p, q)dx\,dy, \tag{19}$$

where A_0 is the area bounded by the curve C_0 (see Figure 3), where $p \equiv z_x$, $q \equiv z_y$, and where f is a function of class C''' for all sets (x, y, z, p, q) sufficiently near those on that portion of the surface $z = z(x, y)$ which is bounded by C_0'. We observe that

$$I(a) = \iint_{A_a} (x, y, z+a\zeta, p+a\zeta_x, q+a\zeta_y)dx\,dy \tag{20}$$

is the value of the integral (19) taken over the portion of the surface (17) which is bounded by its intersection with the surface $\varphi(x, y, z) = 0$. Our problem is to obtain the first and second derivatives $I'(0)$ and $I''(0)$ of this integral in symmetric, usable forms. In doing so, we shall assume that $f \neq 0$ on the surface $z = z(x, y)$ along its intersection C_0' with the surface $\varphi = 0$. This is a customary assumption for problems of the calculus of variations with variable limits, the reason for which will appear later.

3. **The first variation.** We apply now to the integral (20)

$$I(a) = \iint_{A_a} f(x, y, z+a\zeta, p+a\zeta_x, q+a\zeta_y)dx\,dy$$

the result in the first of equations (9). Replacing g in that equation by f, we obtain

$$I'(0) = \iint_{A_0} f_a\,dx\,dy + \int_0^l f v_a\,du, \tag{21}$$

where A_0 is as defined in §2 (see also Figure 3), where

$$f_a = f_z\zeta + f_p\zeta_x + f_q\zeta_y, \tag{22}$$

and where v_a can be computed as follows. Consider the boundary C_a'. It lies on the surface $z = z(x, y) + a\zeta(x, y)$ and on the surface $\varphi(x, y, z) = 0$, and when we transform x, y by equation (3), the function $v(u, a)$, which in equations (18) determines the family of curves C_a, is a solution of the equation

$$\varphi[\xi+v\eta', \eta-v\xi', z(\xi+v\eta', \eta-v\xi')+a\zeta(\xi+v\eta', \eta-v\xi')] = 0, \tag{23}$$

which contains the variables u, v, a. Differentiating (23) with respect to a, we obtain

$$v_a = -\varphi_a/\varphi_v. \tag{24}$$

But when $a=0$,

$$\varphi_a = \varphi_z\zeta, \qquad \varphi_v = (\varphi_x + p\varphi_z)\eta' - (\varphi_y + q\varphi_z)\xi'. \tag{25}$$

Hence

$$v_a = \frac{-\varphi_z\zeta}{(\varphi_x + p\varphi_z)\eta' - (\varphi_y + p\varphi_z)\xi'}. \tag{26}$$

The denominator of the right member of (26) is φ_v. It is different from zero; otherwise we should have the two equations

$$(\varphi_x + p\varphi_z)\eta' - (\varphi_y + q\varphi_z)\xi' = 0,$$
$$(\varphi_x + p\varphi_z)\xi' + (\varphi_y + q\varphi_z)\eta' = 0 \quad \text{(by (29))},$$

and therefore $\varphi_x + p\varphi_z = \varphi_y + q\varphi_z = 0$, since $(\xi', \eta') \neq (0, 0)$. This would imply that $p : q : -1 = -\varphi_x/\varphi_z : -\varphi_y/\varphi_z : -1$ and hence that the two surfaces $z = z(x, y)$ and $\varphi(x, y, z) = 0$ are tangent to each other. But with $f \neq 0$, which we have supposed, this is impossible, as will be seen in §3.

Using (22) and (26) in (21), we now obtain the first derivative

$$I'(0) = \iint_{A_0} (f_z\zeta + f_p\zeta_x + f_q\zeta_y)\,dx\,dy - \int_0^l \frac{\zeta f \varphi_z\,du}{(\varphi_x + p\varphi_z)\eta' - (\varphi_y + q\varphi_z)\xi'}. \tag{27}$$

This result we express in the following theorem:

THEOREM 2. *The first derivative $I'(0)$ of the double integral $I(a)$ of equation* (20), *taken over the portion of the surface $z = z(x, y) + a\zeta(x, y)$ bounded by its intersection with the surface $\varphi(x, y, z) = 0$, has the value given by* (27).

In certain applications, it is desirable to change the form of (27) by performing an integration by parts on the double integral and then applying to the result Green's theorem for the plane. Thus we may write

$$f_p\zeta_x = \frac{\partial}{\partial x} f_p\zeta - \zeta\frac{\partial}{\partial x} f_p, \quad f_q\zeta_y = \frac{\partial}{\partial y} f_q\zeta - \zeta\frac{\partial}{\partial y} f_q;$$

then the double integral in (27) takes the form

$$\iint_{A_0}\left[\zeta\Big(f_z - \frac{\partial}{\partial x}f_p - \frac{\partial}{\partial y}f_q\Big) + \frac{\partial}{\partial x}(f_p\zeta) + \frac{\partial}{\partial y}(f_q\zeta)\right]dx\,dy$$
$$= \iint_{A_0}\zeta\left(f_z - \frac{\partial}{\partial x}f_p - \frac{\partial}{\partial y}f_q\right)dx\,dy + \iint_{A_0}\left[\frac{\partial}{\partial x}(f_p\zeta) + \frac{\partial}{\partial y}(f_q\zeta)\right]dx\,dy.$$

Applying Green's theorem for the plane to the last written integral, we find

$$\iint_{A_0}\left[\frac{\partial}{\partial x}(f_p\zeta)+\frac{\partial}{\partial y}(f_q\zeta)\right]dx\,dy=\int_0^l \zeta(f_p\eta'-f_q\xi')du.$$

Now substituting in (27) the value found for the double integral, we obtain

$$\text{(28)}\quad I'(0)=\iint_{A_0}\zeta\left(f_z-\frac{\partial f_p}{\partial x}-\frac{\partial f_q}{\partial y}\right)dx\,dy$$
$$+\int_0^l \zeta\left[\frac{(f_p\eta'-f_q\xi')((\varphi_x+p\varphi_z)\eta'-(\varphi_y+q\varphi_z)\xi')-\zeta f\varphi_z}{(\varphi_x+p\varphi_z)\eta'-(\varphi_y+q\varphi_z)\xi'}\right]du.$$

Since, along C_0', $\varphi[\xi,\eta,\zeta(\xi,\eta)]\equiv 0$ in u, we can simplify the line integral of (28). Differentiating this identity, we obtain

$$\text{(29)}\quad (\varphi_x+p\varphi_z)\xi'+(\varphi_y+q\varphi_z)\eta'=0\,.$$

Hence from equations (25) and (29), we have

$$\text{(30)}\quad \xi'=-(\varphi_y+q\varphi_z)/\varphi_v,\qquad \eta'=(\varphi_x+p\varphi_z)/\varphi_v,$$

where

$$\text{(31)}\quad \varphi_v^2=(\varphi_x+p\varphi_z)^2+(\varphi_y+q\varphi_z)^2.$$

Substitution from these formulas (30), (31) in the line integral of (28) and application of the relation $\xi'^2+\eta'^2=1$, reduces (28), after simple computation, to the form

$$\text{(32)}\quad I'(0)=\iint_{A_0}\zeta\left(f_z-\frac{\partial}{\partial x}f_p-\frac{\partial}{\partial y}f_q\right)dx\,dy$$
$$+\int_0^l \zeta\left[\frac{f_p(\varphi_x+p\varphi_z)+f_q(\varphi_y+q\varphi_z)-f\varphi_z}{(\varphi_x+p\varphi_z)\eta'-(\varphi_y+q\varphi_z)\xi'}\right]du.$$

We therefore have the following corollary to Theorem 2.

COROLLARY 1. *The first derivative $I'(a)$ of the double integral $I(a)$, of equation* (20), *taken over the portion of the surface $z=z(x,y)+a\zeta(x,y)$ bounded by its intersection with the surface $\varphi(x,y,z)=0$, has the value given by* (32).

In case $z=z(x,y)$ is a minimizing surface for the integral I, $I'(0)$ must vanish. But it is from this result (32) that the Euler necessary condition for double integrals with fixed limits, namely

$$\text{(33)}\quad f_z-\frac{\partial}{\partial x}f_p\zeta-\frac{\partial}{\partial y}f_q\zeta=0,$$

is obtained by the usual proof* and this condition is of course necessary

* For example, see Bolza's *Vorlesungen über Variationsrechnung*, p. 655.

for a minimum in the case of variable limits. When the double integral (32) vanishes, the line integral there must also vanish for every $\zeta(x, y)$. By a slight modification of the fundamental lemma for double integrals, we therefore have further the transversality condition*

$$f_p\varphi_x+f_q\varphi_y+(pf_p+qf_q-f)\varphi_z=0. \tag{34}$$

Hence

COROLLARY 2. *In case $z=z(x, y)$ is a minimizing surface for the double integral*

$$I=\iint_{A_0} f(x, y, z, p, q)dx\,dy,$$

the Euler equation (33) *must hold at every point of the portion of the surface $z=z(x, y)$ inside C_0', and the transversality condition* (34) *must hold at every point of the boundary C_0', which is the line of intersection of the surfaces $z=z(x, y)$ and $\varphi(x, y, z)=0$.*

From the hypothesis made in §2 that $f\neq 0$ along C_0', it follows that the surface $z=z(x, y)$ cannot be tangent to the surface $\varphi(x, y, z)=0$ at any point of their intersection C_0'.

For the special case of the minimal surface, $f=\sqrt{1+p^2+q^2}$ and the transversality condition (34) reduces to

$$p\varphi_x+q\varphi_y-\varphi_z=0,$$

which shows that the surfaces $z=z(x, y)$ and $\varphi(x, y, z)=0$ are orthogonal to each other.

4. **The second variation.** To get $I''(0)$, we apply to the integral (20),

$$I(a)=\iint_{A_a} f(x, y, z+a\zeta, p+a\zeta_x, q+a\zeta_y)dx\,dy,$$

the result obtained in the second of equations (9), §1. Replacing g in that equation by f, we obtain

$$I''(0)=\iint_{A_0} f_{aa}\,dx\,dy+\int_0^l L\,du, \tag{35}$$

where $L=f(v_{aa}+v_a^2/\rho)+f_v v_a^2+2f_a v_a$. The function v_a of course has the value given by (26), and v_{aa} can be computed by differentiating equation (24), $\varphi_v v_a+\varphi_a=0$, with respect to a. The result of this differentiation is

$$v_{aa}=-\frac{1}{\varphi_v}(\varphi_{vv}v_a^2+2\varphi_{av}v_a+\varphi_{aa}). \tag{36}$$

---

* See Bolza, loc. cit., p. 671.

Carrying out the differentiation indicated in (36) and remembering that the arguments in φ are

$$\xi+v\eta', \quad \eta-v\xi', \quad z(\xi+v\eta', \eta-v\xi')+a\zeta(\xi+v\eta', \eta-v\xi'),$$

we obtain after collecting terms suitably

$$(37) \quad v_{aa}=-\frac{1}{\varphi_v^3}\Big\{\varphi_z^2\zeta^2\big[\varphi_{xx}\eta'^2+\varphi_{yy}\xi'^2+\varphi_{zz}(p\eta'-q\xi')^2-2\varphi_{xy}\xi'\eta'$$
$$+2\varphi_{xz}\eta'(p\eta'-q\xi')-2\varphi_{yz}\xi'(p\eta'-q\xi')+\varphi_z(r\eta'^2-2s\xi'\eta'+t\xi'^2)\big]$$
$$-2\varphi_z\zeta\varphi_v\big[\varphi_{xz}\eta'\zeta-\varphi_{yz}\xi'\zeta+\varphi_{zz}(p\eta'-q\xi')\zeta+\varphi_z(\zeta_x\eta'-\zeta_y\xi')\big]+\big[\varphi_{zz}\zeta^2\varphi_v^2\big]\Big\},$$

where, according to custom, r, s, t stand for the derivatives $r=z_{xx}$, $s=z_{xy}$, $t=z_{yy}$, and where $\varphi_v\neq 0$, as was explained just after equation (26). But on collecting the coefficients of ξ'^2, η'^2, and $\xi'\eta'$ in (37), we find for v_{aa} the value

$$(38) \quad v_{aa}=\frac{\zeta^2}{\Delta^3}\big[(r_1-r)\eta'^2-2(s_1-s)\xi'\eta'+(t_1-t)\xi'^2\big]+\frac{2\zeta}{\Delta^3}\big[\zeta_x(p-p_1)+\zeta_y(q-q_1)\big],$$

where

$$\Delta=\sqrt{(p-p_1)^2+(q-q_1)^2}, \quad p_1=-\varphi_x/\varphi_z, \quad q_1=-\varphi_y/\varphi_z,$$

$$(39) \quad \begin{aligned} r_1&=\partial p_1/\partial x=-\frac{1}{\varphi_z^3}(\varphi_{xx}\varphi_z^2-2\varphi_x\varphi_z\varphi_{xz}+\varphi_x^2\varphi_{zz}),\\ s_1&=\partial q_1/\partial x=-\frac{1}{\varphi_z^3}(\varphi_{xy}\varphi_z^2-\varphi_y\varphi_z\varphi_{xz}-\varphi_x\varphi_z\varphi_{yz}+\varphi_x\varphi_y\varphi_{zz}),\\ t_1&=\partial q_1/\partial y=-\frac{1}{\varphi_z^3}(\varphi_{yy}\varphi_z^2-2\varphi_y\varphi_z\varphi_{yz}+\varphi_y^2\varphi_{zz}).\end{aligned}$$

The introduction of p_1, q_1, r_1, s_1, t_1, thus putting φ_z^3 in the denominators of p_1, q_1, r_1, s_1, t_1, seems to require that the surface $\varphi(x, y, z)=0$ shall be representable in the form $z=z_1(x, y)$, but in the final form of the integrand of the line integral to be obtained, we shall see that this apparent requirement is not essential. The other three terms of L are fv_a^2/ρ and the two terms

$$(40) \quad \begin{aligned} f_v v_a^2&=\frac{\zeta^2}{\Delta^3}\big[(f)_x(p-p_1)+(f)_y(q-q_1)\big],\\ 2f_a v_a&=-\frac{2\zeta^2}{\Delta}f_z-\frac{2\zeta}{\Delta}(f_p\zeta_x+f_q\zeta_y),\end{aligned}$$

where

$$(41) \quad \begin{aligned} (f)_x&=f_x+pf_z+rf_p+sf_q,\\ (f)_y&=f_y+qf_z+sf_p+tf_q.\end{aligned}$$

The terms of L may therefore be collected so as to give the sum E_1+E_2, where

$$
(42)\qquad
\begin{aligned}
E_1 &\equiv \frac{\zeta^2}{\Delta^3}\Big[f(r_1-r)\eta'^2-2f(s_1-s)\xi'\eta'+f(t_1-t)\xi'^2+\frac{f\Delta}{\rho}+(f)_z(p-p_1)\\
&\qquad +(f)_y(q-q_1)-2f_z\Delta^2\Big],\\
E_2 &\equiv \frac{2\zeta}{\Delta^3}\big[f(p-p_1)\zeta_x+f(q-q_1)\zeta_y-(f_p\zeta_x+f_q\zeta_y)\Delta^2\big].
\end{aligned}
$$

At first thought one might think that the expression E_1+E_2, copied from (42), is the simplest form obtainable for L. We next show that this is not the case. In fact, we shall eliminate from L the derivatives ζ_x and ζ_y, which occur only in E_2.

In E_2 use $f=f_p(p-p_1)+f_q(q-q_1)$, an immediate consequence of the transversality condition (34). Then, upon multiplying out the products in E_2 and re-collecting its terms suitably, we find

$$
(43)\qquad E_2=2(f_p\xi'+f_q\eta')\,(\zeta_x\xi'+\zeta_y\eta')v_a,
$$

where we are to observe from (30) and (26) that

$$
\xi'=-(q-q_1)/\Delta,\qquad \eta'=(p-p_1)/\Delta,\qquad v_a=-\zeta/\Delta.
$$

Fortunately, the factor $\zeta_x\xi'+\zeta_y\eta'$ of E_2, in (43), occurs in the formula for v_a' found from the solution $v(u,a)$ of $\varphi=0$ (see equation (23)). Indeed one readily finds that the equation defining v_a' is

$$
(44)\qquad \zeta_x\xi'+\zeta_y\eta'=\big[(r_1-r)-(t_1-t)\big]\xi'\eta'v_a+(s_1-s)\,(\eta'^2-\xi'^2)v_a-v_a'\Delta
$$

after neglecting a term which has $\xi'\xi''+\eta'\eta''=0$ as a factor. Using in (43) the value of $\zeta_x\xi'+\zeta_y\eta'$ which (44) gives, we obtain a second new form of E_2, namely,

$$
(45)\qquad
\begin{aligned}
E_2=2(f_p\xi'+f_q\eta')v_a\big[(r_1-r)-(t_1-t)\big]\xi'\eta'v_a+(s_1-s)\,(\eta'^2-\xi'^2)v_a\\
-2v_av_a'\Delta(f_p\xi'+f_q\eta').
\end{aligned}
$$

The only part of E_2 which involves ζ_x and ζ_y is v_a', which we eliminate presently. Let $S\equiv\Delta(f_p\xi'+f_q\eta')$, so that the last term of E_2 in (45) is $-S\cdot 2v_av_a'$. Since this term is in L, let us integrate it by parts. We have

$$
(46)\qquad -\int_0^l S\cdot 2v_av_a'\,du=-Sv_a^2\Big]_0^l+\int_0^l v_a^2S'\,du,
$$

where the first term on the right in (46) vanishes since Sv_a^2 has the period l, and where, after using the relations $\xi''=-\eta'/\rho$, $\eta''=\xi'/\rho$, we find that

$$
(47)\qquad
\begin{aligned}
S'=\big[-f_p\eta'/\rho+f_q\xi'/\rho+\xi'^2(f_p)_x+\eta'^2(f_q)_y+((f_p)_y+(f_q)_x)\xi'\eta'\big]\Delta\\
+\big[(r-r_1)-(t-t_1)\big]\xi'\eta'+(s-s_1)\,(\eta'^2-\xi'^2).
\end{aligned}
$$

The integration by parts thus gives us the result that we may replace the last term of E_2 in (45) by $v_a^2 S'$, where S' is defined by (47). Hence we may replace E_2 in L by

$$
\begin{aligned}
(48)\quad E=&(f_p\xi'+f_q\eta')v_a[(r_1-r)-(t_1-t)]\xi'\eta'v_a+(s_1-s)(\eta'^2-\xi'^2)v_a\\
&+v_a^2\Delta[(f_p)_x\xi'^2+(f_q)_y\eta'^2+((f_p)_y+(f_q)_x)\xi'\eta']+v_a^2\Delta(-f_p\eta'/\rho+f_q\xi'/\rho),
\end{aligned}
$$

where $v_a^2\Delta(-f_p\eta'+f_q\xi')/\rho=-fv_a^2/\rho$, as is seen immediately from formulas (30) and (34). But $-fv_a^2/\rho$ exactly cancels with $f\zeta^2\Delta/\rho\Delta^3=fv_a^2/\rho$ in E_1 (see equation (42)). By collecting the remaining terms of E_1 and E, after using the values of ξ', η', and v_a written just below equation (43), and after putting $f=f_p(p-p_1)+f_q(q-q_1)$ in E_1, and also after neglecting a term equal to the Euler expression (see equation (33)), we obtain the following form for $I''(0)$:

$$
(49)\qquad I''(0)=\iint_{A_0}2\,\Omega\,dx\,dy+\int_0^l\zeta^2(A_1+A_2)du,
$$

where

$$
\begin{aligned}
&f_{aa}\equiv 2\Omega\equiv f_{zz}\zeta^2+2f_{pz}\zeta\zeta_x+f_{pp}\zeta_x^2+2f_{pq}\zeta_x\zeta_y+2f_{qz}\zeta\zeta_y+f_{qq}\zeta_y^2,\\
&A_1\equiv\frac{1}{\Delta^3}[f_p(p-p_1)r_1+(f_p(q-q_1)+f_q(p-p_1))s_1+f_q(q-q_1)t_1],\\
(50)\quad &A_2\equiv-\frac{1}{\Delta^3}\Big\{[f_p(p-p_1)r+(f_p(q-q_1)+f_q(p-p_1))s+f_q(q-q_1)t]\\
&\qquad\qquad-[(f)_x(p-p_1)+(f)_y(q-q_1)]\\
&\quad+[(f_p)_y+(f_q)_x](p-p_1)(q-q_1)+(f_p)_x(p-p_1)^2+(f_q)_y(q-q_1)^2\\
&\quad+[f_z+(f_p)_x](p-p_1)^2+[f_z+(f_q)_y](q-q_1)^2\Big\},
\end{aligned}
$$

the meaning of the notations $(f_p)_x$, $(f_q)_x$, etc. being obvious from equations (41).

Two things are to be noted here. First, the only place where the curvature of the φ-surface appears in the line integral of (49) is in A_1, which contains linearly the elements r_1, s_1, t_1, and also p, q and p_1, q_1, which define the normals to the z- and φ-surfaces; while A_2 involves the curvature of the z-surface, r, s, t appearing linearly, and the elements p, q and p_1, q_1. Secondly, if in A_1, A_2 we substitute for p_1, q_1, r_1, s_1, t_1 their values given in equations (39), we observe that φ_z does not occur in the denominator of either of the resulting expressions for A_1 and A_2, in (50). Hence we see that the apparent necessity of making the restriction that $\varphi_z\neq 0$ just after equations (39) was not essential.

Equation (49) states the result which we set out to obtain. We express it in the following theorem.

THEOREM 3. *The second derivative $I''(0)$ of the double integral $I(a)$ of equation* (20), *taken over the portion of the surface $z=z(x, y)+a\zeta(x, y)$ bounded by its intersection with the surface $\varphi(x, y, z)=0$, has the value given by the formula* (49), *where f_{aa}, A_1, and A_2 are the expressions defined in equations* (50).

5. **Boundary value problem associated with the second variation.** A new necessary condition in order that the surface $z=z(x, y)$ shall minimize the double integral (19) is easily deducible from equation (49). To obtain this condition, we first perform an integration by parts on the double integral of (49) and then apply Green's theorem to the result. Then according to Euler's theorem on homogeneous functions, we may write

$$(51)\quad \iint_{A_0} 2\,\Omega dx\,dy=\iint_{A_0}\left[\zeta\left(\Omega_\zeta-\frac{\partial}{\partial x}\Omega_{\zeta_x}-\frac{\partial}{\partial y}\Omega_{\zeta_y}\right)+\frac{\partial}{\partial x}(\Omega_{\zeta_x}\zeta)+\frac{\partial}{\partial y}(\Omega_{\zeta_y}\zeta)\right]dx\,dy.$$

Now applying Green's theorem for the plane to the last two terms of the right member of (51), we obtain

$$(52)\quad \iint_{A_0} 2\,\Omega dx\,dy=\iint_{A_0}\zeta\left(\Omega_\zeta-\frac{\partial}{\partial x}\Omega_{\zeta_x}-\frac{\partial}{\partial y}\Omega_{\zeta_y}\right)dx\,dy + \int_0^l \zeta(\Omega_{\zeta_x}\eta'-\Omega_{\zeta_y}\xi')\,du,$$

since u is the arc on C_0. Taking the three equations (49), (50), and (52) into account, we may now write

$$(53)\quad I''(0)=\iint_{A_0}\zeta\Psi(\zeta)\,dx\,dy+\int_0^l \zeta(P\zeta+Q\zeta_x+R\zeta_y)\,du\,,$$

where

$$(54)\quad \begin{aligned}
\Psi(\zeta)&\equiv\Omega_\zeta-\frac{\partial}{\partial x}\Omega_{\zeta_x}-\frac{\partial}{\partial y}\Omega_{\zeta_y}\,,\\
P&\equiv A_1+A_2+[f_{pz}(p-p_1)+f_{qz}(q-q_1)]\frac{1}{\Delta}\,,\\
Q&\equiv[f_{pp}(p-p_1)+f_{pq}(q-q_1)]\frac{1}{\Delta}\,,\\
R&\equiv[f_{pq}(p-p_1)+f_{qq}(q-q_1)]\frac{1}{\Delta}\,,
\end{aligned}$$

Δ being defined in equation (39).

From (53) we can now state the new necessary condition in order that the surface $z=z(x, y)$ shall minimize the double integral (19):

THEOREM 4. *In order that the surface $z=z(x, y)$ shall minimize the double integral* (19) *it is necessary that for negative values of* λ *the boundary value problem*

$$\begin{aligned} &\Psi(\zeta)-\lambda\zeta=0 && \textit{in the region } A_0, \\ &P\zeta+Q\zeta_x+R\zeta_y=0 && \textit{on the arc } C_0 \end{aligned} \tag{55}$$

have no solution except $\zeta\equiv 0$, *the functions* P, Q, R *being defined in equations* (54).

Theorem 4 may be regarded as an analogue of Hilbert's statement of the Jacobi condition for the simplest problem of the calculus of variations in the plane.

The proof that the boundary value problem (55) has no solution except $\zeta\equiv 0$ is seen by observing that the line integral is 0 for every ζ that satisfies the second condition of (55) and that a ζ which satisfies both conditions of (55) gives the double integral in equation (53) the value

$$\lambda\iint_{A_0}\zeta^2\,dx\,dy\,.$$

Now if $\lambda<0$, $I''(0)<0$, since the double integral is positive, and hence the second variation $=\epsilon^2 I''(0)$ is negative. Therefore the double integral I of equation (19) is not minimized.

6. **The minimal surface.** If the surface $z=z(x, y)$ is a surface of minimal area, then the integrand of I is the function $f=\sqrt{1+p^2+q^2}$. The only derivatives of f which are not zero and which appear in P, Q, R in this case are

$$\begin{aligned} &f_p=p/f,\ f_q=q/f,\ f_{pp}=(1+q^2)/f^3,\ f_{pq}=-pq/f^3,\ f_{qq}=(1+p^2)/f^3, \\ &(f)_x=(pr+qs)/f,\quad (f)_y=(ps+qt)/f, \\ &(f_p)_x=[(1+q^2)r-pqs]/f^3,\quad (f_p)_y=[(1+q^2)s-pqt]/f^3, \\ &(f_q)_x=[-pqr+(1+p^2)s]/f^3,\quad (f_q)_y=[-pqs+(1+p^2)t]/f^3. \end{aligned} \tag{56}$$

Using these equations (56) we therefore find for P, Q, R the values

$$\begin{aligned} P\equiv{}&[r_1p(p-p_1)+s_1(2pq-pq_1-p_1q)+t_1q(q-q_1)]\,\frac{1}{f\Delta^3} \\ &-\big[r\big(2(1+q^2)(p-p_1)^2-pq(p-p_1)(q-q_1)\big)+t\big(2(1+p^2)(q-q_1)^2 \\ &-pq(p-p_1)(q-q_1)\big)+s\big((2+p^2+q^2)(p-p_1)(q-q_1)-2pq\Delta^2\big)\big]\frac{1}{f^3\Delta^3}, \\ Q\equiv{}&[(1+q^2)\,(p-p_1)-pq\,(q-q_1)]\,\frac{1}{f^3\Delta}, \\ R\equiv{}&[-pq\,(p-p_1)+(1+p^2)\,(q-q_1)]\frac{1}{f^3\Delta}. \end{aligned} \tag{57}$$

Since $f_{pz}=f_{qz}=0$, in this case, P has the value which A_1+A_2 assumes for the minimal surface. Hence we have the following corollary to Theorem 3.

COROLLARY. *For the case of the minimal surface, the integral* $I''(0)$ *of equation* (49) *takes the form*

$$I''(0)=\iint_{A_0}[(1+q^2)\zeta_x^2-2pq\zeta_x\zeta_y+(1+p^2)\zeta_y^2]\frac{1}{f^3}\,dx\,dy+\int_0^l P\zeta^2du\,,$$

where, in this case, $P=A_1+A_2$ *is defined in equation* (57).

Since we also have special values Q, R (see equations (54)) in this case, we may state the following corollary to Theorem 4.

COROLLARY. *In order that the surface of minimum area shall minimize the double integral* (19) *it is necessary that for negative values of* λ *the boundary value problem*

$$-\frac{\partial}{\partial x}\Omega_{\zeta_x}-\frac{\partial}{\partial y}\Omega_{\zeta_y}-\lambda\zeta=0 \qquad \textit{in the region } A_0\,,$$

$$P\zeta+Q\zeta_x+R\zeta_y=0 \qquad \textit{on the arc } C_0$$

have no solution except $\zeta\equiv 0$, *where the functions* P, Q, R *are now defined by equations* (57) *and where*

$$\Omega_{\zeta_x}=\frac{(1+q^2)\zeta_x-pq\zeta_y}{(1+p^2+q^2)^{3/2}}\,,\quad \Omega_{\zeta_y}=\frac{-pq\zeta_x+(1+p^2)\zeta_y}{(1+p^2+q^2)^{3/2}}\,.$$

UNIVERSITY OF CHICAGO,
CHICAGO, ILL.

ON EXTENDING A CONTINUOUS (1-1) CORRESPONDENCE OF TWO PLANE CONTINUOUS CURVES TO A CORRESPONDENCE OF THEIR PLANES*

BY

HARRY MERRILL GEHMAN

Various authors have considered the following problem: given two sets of points, M and M', lying in planes S and S' respectively, and a continuous (1-1) correspondence† T, such that $T(M)=M'$, under what conditions can the correspondence be extended to the planes? That is, under what conditions does there exist a continuous (1-1) correspondence U, such that $U(S)=S'$, and such that for points of M, U is identical with T?

A. Schoenflies‡ has shown that in case M is a simple closed curve the correspondence can be extended to the planes, without any conditions being imposed.

R. L. Moore§ has shown that if M and M' are subsets of arcs, the correspondence can always be extended to the planes. If we consider only the case where M is a connected set, Moore's theorem applies only to the case where M is an arc.

Moore and Schoenflies, then, have proved that if M is an arc or a simple closed curve, the correspondence can be extended to the planes, without any conditions being imposed on the correspondence. In this paper, we show that if M is any plane continuous curve,‖ the correspondence can

* Presented to the Society, May 3, 1924, and February 28, 1925; received by the editors in March, 1925.

† A correspondence T which sends M into $T(M)$ is said to be *continuous*, if in case the point P of M is a limit point of N, a subset of M, then $T(P)$ is a limit point of $T(N)$. See R. L. Moore, *Report on continuous curves from the viewpoint of analysis situs*, Bulletin of the American Mathematical Society, vol. 29 (1923), p. 289. We shall refer to this paper hereafter as "Report."

‡ A. Schoenflies, *Beiträge zur Theorie der Punktmengen*, Mathematische Annalen, vol. 62 (1906), p. 324. See also J. R. Kline, *A new proof of a theorem due to Schoenflies*, Proceedings of the National Academy of Sciences, vol. 6 (1920), p. 529. A *simple closed curve* is a set which is in continuous (1-1) correspondence with a circle.

§ R. L. Moore, *Conditions under which one of two given closed linear point sets may be thrown into the other one by a continuous transformation of a plane into itself*, American Journal of Mathematics, vol. 48 (1926), p. 67. An *arc* is a set which is in continuous (1-1) correspondence with an interval of a straight line.

‖ For the various definitions of a continuous curve see Report, pp. 289–295.

be extended to the planes, provided that we impose the condition that sides of arcs be preserved under the correspondence. Our theorem includes Schoenflies's theorem as a special case, since our condition is evidently satisfied in this special case.

The following example shows the necessity for imposing some condition on the correspondence T. The continuous curve M in the XY plane consists of the straight line intervals from (0, 0) to (3, 0), from (1, 0) to (1, 1) and from (2, 0) to (2, 1). The continuous curve M' in the $X'Y'$ plane is given by subjecting the points of M to the following transformation: if $x\neq 2$, $x=x'$ and $y=y'$; if $x=2$, $x=x'$ and $y=-y'$. Here T is a continuous (1-1) correspondence and $T(M)=M'$, but the correspondence evidently cannot be extended to the entire planes.

Definition. If M and M' are continuous curves lying in planes S and S' respectively, and T is a continuous (1-1) correspondence such that $T(M)=M'$, we say that *sides are preserved under* T, if, given any arc AB of M, and any simple closed curve J in S containing AB as a subset, then there exists a simple closed curve J' in S' containing $T(AB)=A'B'$ as a subset, and such that if N designates the points of M interior to J, then the interior of J' contains $T(N)=N'$; and also, if given any simple closed curve J_1' in S' containing $A'B'$ as a subset, then there exists a simple closed curve J_1 in S containing AB as a subset, and such that if N_1' designates the points of M' interior to J_1', then the interior of J_1 contains $T^{-1}(N_1')=N_1$.

In the following we shall frequently use this notation: if X is any subset of M, we shall denote $T(X)$ by X'; if Y' is any subset of M', we shall denote $T^{-1}(Y')$ by Y.

Theorem I. *If M and M' are continuous curves containing no simple closed curve* and lying in planes S and S' respectively, and if there exists a continuous* (1-1) *correspondence T, such that $T(M)=M'$, and such that sides are preserved under T, then there exists a continuous* (1-1) *correspondence U, such that $U(S)=S'$, and such that if, for any point P of M, $T(P)=P'$, then $U(P)=P'$.*

Before proceeding with the proof of Theorem I, we shall discuss the definition of "sides preserved under T" for the case where M is a continuous

* For a discussion of this type of continuous curve, see S. Mazurkiewicz, *Un théorème sur les lignes de Jordan*, Fundamenta Mathematicae, vol. 2 (1921), p. 119; R. L. Wilder, abstracts in the Bulletin of the American Mathematical Society, vol. 29 (1923), p. 118, and *Concerning continuous curves*, Fundamenta Mathematicae, vol. 7 (1925), p. 340; and R. L. Moore, *Concerning the cut-points of continuous curves*, Proceedings of the National Academy of Sciences, vol. 9 (1923), p. 101.

curve containing no simple closed curve. In this discussion we have need of the following lemma.

LEMMA A. *If I_1 and I_2 are two simply connected domains, whose boundaries are B_1 and B_2, and whose outer boundaries* C_1 and C_2 have in common an arc UV such that no point of $\underline{UV}$† is a limit point of $B_1+B_2-\underline{UV}$, then $\underline{UV}$ is in the boundary of two different domains complementary to B_1+B_2, such that either* (1) *one domain is a subset of I_1, the other of I_2, or* (2) *one domain is a subset of both I_1 and I_2, and the other has no points in common with either I_1 or I_2.*

Proof. About each point X of $\underline{UV}$, let us construct a circle C_X whose exterior contains $B_1+B_2-\underline{UV}$. Then corresponding to each point X, we can construct a simple closed curve J_X, formed of an arc X_1XX_3 of $\underline{UV}$ and an arc $\underline{X_1X_2X_3}$ in I_1, and whose interior is in I_1 and the interior of C_X.‡ The sum of the interiors of the simple closed curves J_X is a domain D, because if the boundaries of any pair have an arc of $\underline{UV}$ in common, their interiors have a point in common. The domain D contains no points of B_1+B_2 by construction. If we add to D all points which can be joined to a point of D by an arc having no points in common with B_1+B_2, we obtain a domain D_1 complementary to B_1+B_2, and such that $\underline{UV}$ forms part of the boundary of D_1. Evidently D_1 is a subset of I_1.

By a construction similar to the above, but taking in this case the arc $\underline{X_1X_2X_3}$ exterior to I_1, we obtain a domain D_2 complementary to B_1+B_2, whose boundary contains $\underline{UV}$, and which has no points in common with I_1.

Since $\underline{UV}$ is part of the boundary of I_2, either D_1 or D_2 contains a point of I_2. Since a point of I_2 cannot be joined to a point exterior to I_2 by an arc having no points in common with B_2, it is evident that either D_1 or D_2 is a subset of I_2, and the other has no points in common with I_2. We have accordingly, the two possibilities mentioned in Lemma A.

We shall now consider some consequences of the definition of "sides are preserved under T" for a continuous curve M containing no simple closed curve. Let AB be a maximal arc§ in M. Let J be a simple closed curve in the plane S, containing AB as a subset and containing no other points of M.

---

* R. L. Moore, *Concerning continuous curves in the plane*, Mathematische Zeitschrift, vol. 15 (1922), p. 254.

† If UV is an arc, $\underline{UV}$ denotes $UV-U-V$.

‡ R. L. Moore, *On the foundations of plane analysis situs*, these Transactions, vol. 17 (1916) p. 131. See especially Theorem 28.

§ A *maximal arc* in a set M is an arc which is not a proper subset of any other arc in M. See S. Mazurkiewicz, loc. cit., Lemma 13, p. 129.

Let N be the points of M interior to J. Since sides are preserved under T, there exists in S' a simple closed curve J', containing $A'B'$ as a subset, and enclosing N'. We shall now show that in case J' contains or encloses any points of $M'-(A'B'+N')$, J' can be replaced by a simple closed curve J'' which contains $A'B'$ as a subset, and which encloses N', but which neither contains nor encloses any other points of M'.

We shall first show that A' and B' are on the boundary of the same domain of $(I'+M')-M'$, where I' designates the interior of J'. For, if not, there is an arc $C'D'$ in M', such that (a) C' is on $J'-A'B'$; (b) D' is on $\underset{\smile}{A'B'}$; (c) $\underset{\smile}{C'D'}$ is in I'. Since N' is entirely in I', C' is not a point of N'. Since $N+AB$ is closed, $N'+A'B'$ is closed, and therefore no point of $\underset{\smile}{C'D'}$ can be in the set N'.

Now let J' be the J_1' of the definition. The curve J' encloses N' and $\underset{\smile}{C'D'}$. The corresponding simple closed curve J_1 in S encloses N and $\underset{\smile}{CD}$, and contains the arc AB as a subset. The simple closed curves J and J_1 in S have AB in common, and therefore satisfy the conditions of Lemma A. Since their interiors have N in common, and since limit points of N are on $\underset{\smile}{AB}$, it follows that one of the domains (complementary to $J+J_1$), whose boundary contains $\underset{\smile}{AB}$, consists entirely of points common to the interiors of J and J_1. Since D is a point of $\underset{\smile}{AB}$, and since $\underset{\smile}{CD}$ is interior to J_1, it follows that at least part of $\underset{\smile}{CD}$ is interior to J and therefore in the set N. In that case, the corresponding part of $\underset{\smile}{C'D'}$ lies in N', contrary to a previous statement. Therefore A' and B' are on the boundary of the same connected domain of $(I'+M')-M'$.

The points A' and B' can therefore be connected by an arc* in $I'-M'$, which forms with the arc $A'B'$ of M' a simple closed curve J''. The arc $J''-\underset{\smile}{A'B'}$ separates I' into two parts, and the part enclosed by J'' contains all of N'. The supposition that J'' encloses other points of M' in addition to N' leads to a contradiction similar to that obtained above.

Therefore in case M contains no simple closed curve, and AB is a maximal arc of M, and J contains AB but no other points of M, then we can add to our definition of "sides are preserved under T," that J' contains $A'B'$ but no other points of M', and that the interior of J' contains N' and no other points of M'; and similarly for J_1' and J_1.

Lemma A and the previous discussion show also that if any two simple closed curves have a maximal arc AB of M in common and contain no other points of M, then either their interiors contain the same subset of M, or

---

* Report, pp. 290–291.

their interiors have no points of M in common, in which case, since $M-A-B$ is connected, they contain all of M, save the arc AB.

DEFINITION. If M is a continuous curve containing no simple closed curve, and N is a closed and connected subset of M, we shall call a maximal connected subset of $M-N$ a *tree with respect to* N, or a *tree in* $M-N$. A tree has one and only one limit point in N, which point we shall call the *foot* of the tree. A tree plus its foot forms a closed set.

R. L. Wilder* has proved that the number of trees is countable, and that given any positive number ϵ, there are at most a finite number of trees of diameter greater than ϵ.

Proof of Theorem I. Since M and M' are bounded we can construct in the plane S a circle C containing M in its interior I, and in the plane S' a circle C' containing M' in its interior I'. If AB is a maximal arc of M, we can join A to any point D of C by an arc in $(I-M)+A+D$, and we can join B to any other point E of C by an arc in $(I-M)+B+E-AD$. The arc $A'B'$ in M' is also a maximal arc, and if we select any arbitrary points D' and E' of C', we can join A' to D' and B' to E' by arcs in $(I'-M')+A'+D'$ and $(I'-M')+B'+E'-A'D'$, respectively.

If X and Y are points of C separating D and E, the arcs EXD (of C), DA, AB (of M), and BE form a simple closed curve J, and the arcs EYD (of C), DA, AB and BE form a simple closed curve J_1. The interiors of J and J_1 have no points in common, and the sum of their interiors contains all points of M save AB.

If X' and Y' are any two points of C' separating D' and E', there exist likewise in S' two simple closed curves $J'=E'X'D'A'B'E'$ and $J_1'=E'Y'D'A'B'E'$, whose interiors have no points in common, and the sum of whose interiors contains all points of M' save $A'B'$.

In our previous discussion of "sides are preserved under T," we have shown that under the above conditions, one of the simple closed curves J' or J_1' (suppose the former) will enclose all the points of M' which correspond under T to the points of M which J encloses, and the other, J_1', will enclose the points of M' which correspond to the points of M which J_1 encloses.

Let us select an arbitrary positive number ϵ. Suppose either $M-AB$ or $M'-A'B'$ contains a tree of diameter greater than $\frac{1}{9}\epsilon$. Let T be a tree in $M-AB$ which is interior to J, and let T' be the corresponding tree in $M'-A'B'$ which is interior to J', these trees being such that the diameter of either T or T' is greater than $\frac{1}{9}\epsilon$. Let the foot of T be F, and let FG be

* R. L. Wilder, loc. cit., first paper, Theorem II.

a maximal arc of $(T+F)$ such that the diameter of either FG or $F'G'$ is greater than $\frac{1}{18}\epsilon$. If H is any point of $\underline{DXE}$, H can be joined to G by an arc interior to J, save for H, and having only G in common with M. If H' is any point of $\underline{D'X'E'}$ a similar arc $H'G'$ can be drawn. The arc FGH separates the interior of J into two domains, and the arc $F'G'H'$ similarly separates the interior of J'. It is evident that the points of M interior to the simple closed curve $FGHDAF$ have their corresponding points in M' interior to the simple closed curve $F'G'H'D'A'F'$. Similarly for points interior to $FGHEBF$.

If $M-(AB+FG)$ or $M'-(A'B'+F'G')$ contains a tree of diameter greater than $\frac{1}{9}\epsilon$, the simple closed curves enclosing it and the corresponding tree in the other set can both be separated by arcs in the manner indicated above. After a finite number of steps $M-(AB+FG+\cdots)$ and $M'-(A'B'+F'G'+\cdots)$ will contain no trees of diameter greater than $\frac{1}{9}\epsilon$, otherwise a theorem* due to R. L. Wilder is contradicted. At this stage, this state of affairs exists: the interior of the circle C is divided into a finite number of domains plus boundary points of these domains, where the domains are such that they are bounded by simple closed curves, each of which consists of a single maximal arc of M, and a single arc having no points in common with M save its end points; the interior of C' is divided into the same number of domains plus boundary points, where the domains are bounded by simple closed curves, each of which consists of a single maximal arc of M' and a single arc having no points in common with M' save its end points; if N represents the set of points of M contained and enclosed by one of the simple closed curves in the plane S, the set $T(N)=N'$ will be contained and enclosed by one of the simple closed curves in the plane S', and this simple closed curve will contain or enclose no points of M' which are not in N'; no tree in N (or N') with respect to the maximal arc of M (or M') belonging to the simple closed curve which contains and encloses N (or N'), is of diameter greater than $\frac{1}{9}\epsilon$.

Let K be any one of the simple closed curves in S; PQ the maximal arc of M on K; N the subset of M contained and enclosed by K. Let K' denote the corresponding simple closed curve in S'. We can select a finite set of points $P_1, P_2, \cdots, P_n$, and $P_1', P_2', \cdots, P_n'$, such that (a) $P_1=P$ and $P_1'=P'$; (b) $P_n=Q$ and $P_n'=Q'$; (c) P_i precedes P_{i+1} on PQ, and P_i' precedes P_{i+1}' on $P'Q'$, for $i=1, 2, \cdots, n-1$; (d) the diameter of each of the arcs P_iP_{i+1} and $P_i'P_{i+1}'$ is less than $\frac{1}{9}\epsilon$; (e) no tree in $N-PQ$ or $N'-P'Q'$ has its foot at any of the points P_i or P_i'; (f) $T(P_i)=P_i'$. Let the set K_i

* R. L. Wilder, loc. cit., second paper, Theorem II.

($i=1, 2, \cdots, n-1$) be P_iP_{i+1} plus all trees in $N-PQ$ with feet on P_iP_{i+1}; and let K_i' be $P_i'P_{i+1}'$ plus all trees in $N'-P'Q'$ with feet on $P_i'P_{i+1}'$. Evidently $T(K_i)=K_i'$, and the diameter of each of the sets $K_1, K_2, \cdots, K_{n-1}$, $K_1', K_2', \cdots, K_{n-1}'$ is less than $\frac{1}{3}\epsilon$.

The point P_1 can be joined to P_2 by an arc in the interior of K, having only its end points in common with M. We shall show that under our given conditions P_1 can be joined to P_2 by such an arc whose diameter is less than $\frac{2}{3}\epsilon$. Suppose an arc P_1RP_2 has been constructed whose diameter is greater than $\frac{2}{3}\epsilon$. This arc forms with P_1P_2 a simple closed curve H which encloses $K_1-P_1P_2$, but encloses no other points of M. There also exists a simple closed curve L, enclosing K_1, and such that every point of L plus its interior is at a distance less than $\frac{1}{12}\epsilon$ from K_1,* and therefore such that the diameter of L is less than $\frac{1}{2}\epsilon$. Since the diameter of the arc P_1RP_2 is greater than $\frac{2}{3}\epsilon$, it must contain points exterior to L. The simple closed curves H and L satisfy the conditions† under which there exists a simple closed curve Q which is a subset of $H+L$, contains the arc P_1P_2 of M, and every point of whose interior is interior to both H and L. The arc from P_1 to P_2 in Q which has only its end points in common with M, has no points exterior to L, and is therefore of diameter less than $\frac{2}{3}\epsilon$; it has no points exterior to H and except for its end points, has no points in common with K_1, and therefore has no points in common with M, and lies entirely in the interior of K. This is an arc satisfying the conditions stated.

The point P_1 can therefore be joined to P_2 by an arc $P_1W_1P_2$ in the interior of K which forms with the arc P_1P_2 of M a simple closed curve C_1, enclosing $K_1-P_1P_2$ but no other points of M, and containing P_1P_2 but no other points of M, and such that the diameter of C_1 is less than ϵ. In the same way an arc $P_2W_2P_3$ can be constructed in the interior of K (save for P_2 and P_3) and exterior to C_1 (save for P_2), which forms with P_2P_3 of M a simple closed curve C_2, enclosing and containing the set K_2 and no other points of M and such that the diameter of C_2 is less than ϵ. In this way we construct the simple closed curves $C_1, C_2, \cdots, C_{n-1}$, all of diameter less than ϵ. In the same way we construct in S' the simple closed curves $C_1', C_2', \cdots, C_{n-1}'$, all of diameter less than ϵ.

Note that if from the interior of K we remove the simple closed curves $C_1+C_2+\cdots+C_{n-1}$ and their interiors, there remains a domain whose

* R. L. Moore, *Concerning the separation of point-sets by curves*, Proceedings of the National Academy of Sciences, vol. 11 (1925), p. 469.

† R. L. Moore, *On the Lie-Riemann-Helmholtz-Hilbert problem of the foundations of geometry*, American Journal of Mathematics, vol. 41 (1919), p. 299. See especially Theorem 26.

boundary is the simple closed curve consisting of the arcs $K-\underbrace{P_1P_n}$, $P_1W_1P_2$, $P_2W_2P_3, \cdots, P_{n-1}W_{n-1}P_n$; similarly for K'.

If the interior of any of the simple closed curves $C_1, \cdots, C_{n-1}$ contains points of M, we shall separate its interior as we have separated the interiors of J and J' into a finite set of domains free from points of M, plus a finite set of simple closed curves (such as C_1) of diameter less than $\frac{1}{2}\epsilon$, the same being true of M'. We shall then continue this process with such of these simple closed curves as enclose points of M.

We shall now define a continuous (1-1) correspondence U, such that $U(S)=S'$, and we shall show that if, for any point P of M, $T(P)=P'$, then $U(P)=P'$.

For points of the circles C and C', U is any continuous (1-1) correspondence between C and C' subject only to these conditions: (1) $U(D)=D'$; $U(E)=E'$; (2) if P is a point of $\underline{EXD}$, $U(P)$ is on $\underline{E'X'D'}$; (3) if P is a point of $\underline{EYD}$, $U(P)$ is on $\underline{E'Y'D'}$; (4) $U(H)=H'$; and similarly for points of C other than H that were joined by arcs to points on trees of M of diameter greater than $\frac{1}{9}\epsilon$, or such that the corresponding trees in M' were of diameter greater than $\frac{1}{9}\epsilon$.

Having defined U for points of the circles C and C', we define U for points exterior to these circles, as being any continuous (1-1) correspondence between S and S' subject only to the condition that for points of C and C' the correspondence be the one defined in the preceding paragraph. That such a correspondence exists, follows from Schoenflies's theorem.

For points interior to C, and on one of the simple closed curves K, consisting of an arc PQ of M, an arc QX interior to C save for X, an arc XY on C, and an arc YP interior to C save for Y, we define U as being any continuous (1-1) correspondence between K and K' subject only to the following conditions: (1) for points of PQ, U is identical with T; (2) for points of XY, U is identical with the correspondence previously defined for points of C.

For points interior to one of the simple closed curves K, and on the arc $P_1W_1P_2W_2P_3\cdots P_{n-1}W_{n-1}P_n$, we define U as being any continuous (1-1) correspondence between $P_1W_1\cdots P_n$ and $P_1'W_1'\cdots P_n'$ subject only to the following conditions: $U(P_1)=P_1'\cdots$, $U(P_n)=P_n'$; and if C_i encloses points of M, the correspondence between $\underbrace{P_iW_iP_{i+1}}$ and $\underbrace{P_i'W_i'P_{i+1}'}$ must be such that $U(H_1)=H_1'$, where H_1 is a point of $\underbrace{P_iW_iP_{i+1}}$ which was joined, by an arc interior to C_i, to a point of a tree, in $M-P_iP_{i+1}$, of diameter greater than $\frac{1}{18}\epsilon$.

For points interior to the simple closed curve formed by $K-\underbrace{P_1P_n}$, and $P_1W_1P_2, \cdots, P_{n-1}W_{n-1}P_n$, U is defined as being any continuous (1-1)

correspondence between S and S' subject only to the condition that for points of the simple closed curves, the correspondence be the one defined in the preceding paragraphs.

If the interior of any of the simple closed curves $C_1, C_2, \cdots, C_n$, say C_1, is free from points of M, U is defined for points of the interior of C_1 as being any continuous (1-1) correspondence between S and S' subject only to the condition that for points of C_1 and C_1' the correspondence be the one already defined for points of C_1 and C_1'.

In case the interior of any of the simple closed curves $C_1, \cdots, C_n$ contains points of M, we have already indicated the method of separation of its interior into a finite number of domains, and we have indicated above how U is defined for every point of S, save for points interior to simple closed curves of diameter less than ϵ, enclosing points of M. We shall now show that by a continuation of the above process, U is defined for such points.

If P is any point of $S-M$, there is some value of k such that $\epsilon/2^k$ is less than the distance from P to a point of M. Therefore at the step where the diameters of the simple closed curves enclosing points of M are less than $\epsilon/2^k$, P will not lie in the interior of such a simple closed curve, and hence U will have been defined for the point P. Similarly for any point in $S'-M'$.

If P is a point of M, and if U has not already been defined for P, then P lies at each step in the interior of a simple closed curve, consisting of an arc of M and an arc in $S-M$. Let the sequence of simple closed curves be $D_1, D_2, \cdots$, where D_{i+1} lies in D_i plus its interior; let the diameter of D_i be less than $\epsilon/2^i$; and let the arc of M on D_i be X_iY_i. Evidently P is the only limit point of this sequence. The corresponding simple closed curves in S' are $D_1', D_2', \cdots$, and D_{i+1}' lies in the interior of D_i', the diameter of D_i' is less than $\epsilon/2^i$, and the arc of M on D_i' is $X_i'Y_i'$, where $X_i'=T(X_i)$, and $Y_i'=T(Y_i)$. The sequence $D_1', D_2', \cdots$ approaches a single point P' as a sequential limit point, and we define $U(P)$ as P'.

It remains to be shown that $T(P)$ is the point $U(P)=P'$. The points $X_1, X_2, \cdots$ of M approach P as a sequential limit. Since T is continuous, the point in M' approached by $X_1', X_2', \cdots$ is $T(P)$, and we have called this point $U(P)=P'$. Therefore U is identical with T for all points of M.

The correspondence U as defined above satisfies the conditions of Theorem I.

Definition. If M and M' are continuous curves lying in planes S and S' respectively, and T is a continuous (1-1) correspondence such that $T(M)=M'$, we say that *interiors are preserved under* T if, given any simple closed curves J of M and J' of M', such that $T(J)=J'$, and if N is the set

of points of M interior to J, and N' is the set of points of M' interior to J', then $T(N)=N'$.

LEMMA B. *If sides are preserved under T, interiors are preserved under T.*

Proof. Given $T(M)=M'$ and sides are preserved under T. Suppose P, a point of M interior to J, is such that $T(P)=P'$ is exterior to J'. We shall show that this leads to a contradiction.

Let PQ be an arc in M, having only Q in common with J, and therefore interior to J, save for Q. Then $P'Q'$ will be exterior to J', save for Q'.

Let A and B be two points of J, $A\neq Q\neq B$, and let D be a point of J, such that D and Q separate A and B. In the exterior of J we shall construct an arc $\underline{AXB}$, such that BQA (of J) plus $\underline{AXB}$ forms a simple closed curve C containing in its interior the interior of J, and therefore $\underline{ADB}$ and $\underline{QP}+P$.

Since sides are preserved under T, there exists a simple closed curve C' in S' containing $B'Q'A'$ and whose interior contains $\underline{A'D'B'}$ and $\underline{Q'P'}+P'$. The curve C' is composed of the arcs $B'Q'A'$ of J', and $\underline{B'X'A'}$ in the exterior of J'. The interior of C' therefore contains the interior of J', in fact $\underline{A'D'B'}$ divides the interior of C' into two parts, the interior of $J'=A'D'B'Q'A'$ and the interior of $A'D'B'X'A'$. Since $\underline{Q'P'}+P'$ is exterior to J' and interior to C', it must lie within $A'D'B'X'A'$. Since T is continuous, the limit point Q' of $\underline{Q'P'}+P'$ must lie on or within $A'D'B'X'A'$, whereas the arc $\underline{B'Q'A'}$ is exterior to $A'D'B'X'A'$. This is the desired contradiction.

THEOREM II. *If M and M' are continuous curves lying in planes S and S' respectively, and if there exists a continuous* (1-1) *correspondence T, such that $T(M)=M'$, and such that sides are preserved under T, then there exists a continuous* (1-1) *correspondence U, such that $U(S)=S'$, and such that if, for any point P of M, $T(P)=P'$, then $U(P)=P'$.*

Proof. Since we have considered in Theorem I the case where M contains no simple closed curve, we shall assume here that M contains at least one simple closed curve. The proof of Theorem II follows, in the main, that of Theorem I. We shall go into detail only when the proof differs from that of Theorem I.

By Schoenflies's definition of a continuous curve,* $S-M$ consists of one unbounded domain and a countable set of bounded domains, the same being true for $S'-M'$. The outer boundary of each bounded domain is a simple closed curve J whose interior contains a countable set of maximal

* Report, p. 290 and p. 295.

connected subsets of $M-J$ (which we shall call "trees"), each having one and only one limit point on the simple closed curve.

If M is a continuous curve, and N is a closed and connected subset of M, we shall call a maximal connected subset of $M-N$ which has one and only one limit point in N a *tree with respect to* N or a *tree in* $M-N$. The limit point in N is called the *foot* of the tree. The number of trees is countable, and if N is the outer boundary of a domain complementary to M, then not more than a finite number of trees can be of diameter greater than any given positive number. In case M contains no simple closed curve this definition is equivalent to the one given previously.

If α is the boundary of D_1, the unbounded domain in $S-M$, then $T(\alpha)=\alpha'$ is the boundary of D_1', the unbounded domain in $S'-M'$. For suppose it were possible that for some point P of α, $T(P)=P'$ is not in β', the boundary of D_1'. Then P' is separated from any point Q' in D_1' by β', and P' and Q' are therefore separated by some simple closed curve J' which is a subset of β'.* The point Q' is not interior to J'; therefore P' is interior to J'. But P is in α and is therefore not interior to $T^{-1}(J')=J$ or any other simple closed curve in M, and in this case interiors (and, therefore, sides) have not been preserved under T, contrary to hypothesis. A similar contradiction is arrived at if we suppose that for some point P' of β', $T^{-1}(P')=P$ is not a point of α. Therefore $\beta'=\alpha'$.

We shall next show that if α is the boundary of D_2, a bounded domain in $S-M$, then $T(\alpha)=\alpha'$ is the boundary of a bounded domain in $S'-M'$, which we may call D_2'. The boundary α contains a simple closed curve J, and α' contains J'. If N denotes the points of M interior to J, N consists of a countable set of trees. Since N' is identical with the points of M' interior to J', any maximal connected subset of N' is a tree with respect to J', otherwise T is not (1-1) and continuous, and therefore N' also consists of a countable set of trees. Let these be $N_1', N_2', N_3', \cdots$. Let L_i' be N_i' plus its foot on J' plus any points of $S'-M'$ lying in a complementary domain of M' whose boundary is in N_i'. The sets $L_1', L_2', L_3', \cdots$ are closed and connected and have no points in common. Furthermore, only a finite number of these sets can be of diameter greater than any given positive number. Under these conditions, $L_1'+L_2'+L_3'+\cdots$ is not connected.† Therefore not every point of the interior of J' lies in $L_1'+L_2'+\cdots$. Let P' be a point of the interior of J' not in $L_1'+L_2'+\cdots$, and let D_2' be the domain in $S'-M'$

---

* R. L. Moore, *Concerning continuous curves in the plane*, loc. cit., Theorem V.

† See abstract by J. R. Kline, Bulletin of the American Mathematical Society, vol. 31 (1925), p. 300.

containing P'. The outer boundary of D_2' is a simple closed curve K'. If K' were entirely interior to J' or had one point in common with J', then K', D_2', and P' would be points of one of the sets L_i', contrary to our selection of P'. If K' had an arc in common with J', but were not identical with J', K' would contain an arc interior to J' joining two points of J', and such an arc cannot exist in any of the trees N_1', N_2', $\cdots$. Therefore, K' is identical with J'. If we suppose that any point P of α is such that P' is not in the boundary of D_2', or that any point Q' of the boundary of D_2' is such that Q is not in α, we obtain a contradiction by the same argument as that used in the case of the boundary of the unbounded domain. Therefore the boundary of D_2' is α'.

We shall next discuss the definition of "sides are preserved under T" for the case where M is any continuous curve.

Let D be any domain in $S-M$, and A and B two points on its boundary β; then any arc AXB such that $\underline{AXB}$ is in D, divides D into two domains, D_1 and D_2, such that $D=D_1+D_2+\underline{AXB}$. The boundary of D_i $(i=1, 2)$ is composed of $\underline{AXB}$ and β_i, a subset of β. Evidently, $\beta=\beta_1+\beta_2$. Suppose that A and B are two points such that β_1 and β_2 are the same for any choice of the arc $\underline{AXB}$. This will be the case if each of the points A and B is a non-cut point of β, i. e., a point whose removal does not disconnect β. If in $\beta_i (i=1, 2)$ we draw the arc AYB (there is only one such arc), and draw in D an arc $\underline{AXB}$ forming a simple closed curve J with AYB, then the discussion under Theorem I of "sides are preserved under T" holds with slight modifications. Therefore, under the above conditions, we can add to our definition of "sides are preserved under T" that J' contains $A'B'$ but no other points of M', that $J'-A'B'$ lies in D' (where D' is the domain in $S'-M'$ whose boundary is $T(\beta)=\beta'$), and that the interior of J' contains N' and no other points of M'.

As in the proof of Theorem I, we can construct a circle C containing M in its interior I, and a circle C' containing M' in its interior I'. Let α denote the boundary of the unbounded domain in $S-M$. Since α is a non-dense continuous curve separating the plane, α contains a simple closed curve J.* Let A and B be two points of J which are not feet of trees in $\alpha-J$. We can join A to D, any point of C, by an arc in $(I-M)+A+D$, and we can join B to any other point E of C by an arc in $(I-M)+B+E-AD$. If D' and E' are arbitrary points on C' we can draw similar arcs $A'D'$ and $B'E'$. Each of the arcs AB of J forms with AD, BE, and the proper one of the arcs DE of C a simple closed curve, and these two simple closed

* Report, pp. 295–296.

curves C_1 and C_2 contain and enclose all points of α. Similarly in S' are two simple closed curves C_1' and C_2' such that the points contained and enclosed by C_i' $(i=1, 2)$ are the points corresponding to those contained and enclosed by C_i.

Let us select an arbitrary positive number ϵ. Let $D_1, D_2, \cdots, D_n$ be a finite set of domains complementary to M, and $D_1', D_2', \cdots, D_n'$ a set of domains complementary to M', such that (1) if α_i is the boundary of D_i $(i=1, 2, \cdots, n)$, then $T(\alpha_i)=\alpha_i'$ is the boundary of D_i' and (2) all domains complementary to either M or M' and of diameter greater than ϵ occur in one set or the other. In each of the bounded domains, such as D_2, select two points A and B on the outer boundary J_2, which are not feet of trees in α_2-J_2, and join A to B by an arc in D_2 save for its end points. If we join A' to B' by an arc in D_2' save for its end points, this arc divides the interior of J_2' into the interiors of two simple closed curves, such that the points of M' contained and enclosed by either one of them are the points corresponding to those contained and enclosed by the corresponding one of the simple closed curves in S which AB forms with J_2.

Let us now consider any one of the above simple closed curves K formed by an arc AXB in M and $\underline{AYB}$ in $S-M$, and its corresponding simple closed curve $K'=A'X'B'Y'A'$ in the plane S'. Let the points of M contained and enclosed by K be denoted by N. Then $T(N)=N'$ is the set contained and enclosed by K'. Suppose either $N-AXB$ or $N'-A'X'B'$ contains a tree of diameter greater than $\frac{1}{9}\epsilon$, and let this tree and its corresponding one be denoted by T and T'. For definiteness let us suppose that the diameter of T is greater than $\frac{1}{9}\epsilon$. Let the foot of T be F. Let G_1 denote some point of T which belongs to α, the boundary of the complementary domain of M in which $\underline{AYB}$ was drawn, and such that an arc FG_1 of α is of diameter greater than $\frac{1}{18}\epsilon$. If $\alpha-G_1$ is connected, we shall denote G_1 by G. If $\alpha-G_1$ is not connected, it consists of a countable set of trees. Let W denote the one which contains F. The set $\alpha-W$ is closed and connected. It has therefore a non-cut point G distinct from G_1. The sets $(\alpha-W-G)$ and $(W+G_1)$ are connected and have G_1 in common. Therefore their sum, $\alpha-G$, is connected, and G is a non-cut point of α. Moreover, any arc in α from F to G passes through G_1 and therefore α contains an arc FG of diameter greater than $\frac{1}{18}\epsilon$.

If H is any point of $\underline{AYB}$, H can be joined to G by an arc interior to K, save for H, and having only G in common with M. In S', a similar arc $H'G'$ can be drawn. The points of M on and interior to $GFAHG$ will have their corresponding points on and interior to $G'F'A'H'G'$. Continuing this process a finite number of times we eventually arrive at the point where none

of the simple closed curves K and K' contain a tree of diameter greater than $\frac{1}{9}\epsilon$.

The choice of sets $K_1, K_2, \cdots, K'_{m-1}$, and the simple closed curves $C_1, C_2, \cdots, C'_{m-1}$ is made as in Theorem I without any modifications. The correspondence U is defined as in Theorem I for all points of $S-M$ and for all points which are boundary points of domains complementary to M. If there are any other points in M, we shall define U as being identical with T.

The correspondence U defined in this way is evidently (1-1), and for points of M, U and T are identical. It remains to be shown that U is continuous, i. e., if $P_1, P_2, \cdots$ is a set of points of S approaching P as a sequential limit, then $P'_1, P'_2, \cdots$ approach $T(P)=P'$ as a sequential limit. In case P is a point of $S-M$, all except a finite number of the points of $P_1+P_2+\cdots$ lie in the same domain as does P, and we have defined U for the complementary domains in such a way that $P'_1, P'_2, \cdots$ will approach P'. In case P is a point of M and an infinite number of points of $P_1+P_2+\cdots$ lie in M, then P' is a limit point of $P'_1+P'_2+\cdots$ because U is identical with T for points of M, and T is continuous. In case P is a point of M, and an infinite number of points of $P_1+P_2+\cdots$ lie in a single domain of $S-M$, then P is a boundary point of that domain, and again U has been defined in such a way that $P'_1, P'_2, \cdots$ have P' as a limit point. In case, finally, P is a point of M, and only a finite number of points of $P_1+P_2+\cdots$ lie in M or in any single domain of $S-M$, then we can pick out an infinite subsequence $Q_1, Q_2, \cdots$, such that each Q_i is a point of $S-M$, and no two points Q_i, Q_j lie in the same domain. Let us associate with each Q_i a point R_i of the boundary of the domain in which Q_i lies. For any given ϵ only a finite number of the points R_i can be selected in such a way that the distance between Q_i and R_i is greater than ϵ. Therefore the set of points $R_1, R_2, \cdots$ also approach P as a sequential limit. Similarly in the plane S', the two sequences $Q'_1, Q'_2, \cdots$ and $R'_1, R'_2, \cdots$ approach the same limit. Since $R'_1, R'_2 \cdots$ are points of M', their limit is P'. Therefore in each case, U is continuous.

The author wishes to express his thanks to Professor John Robert Kline for suggesting this problem and for aiding in its solution.

University of Pennsylvania,
Philadelphia, Pa.

SYSTEMS OF EQUATIONS IN AN INFINITY OF UNKNOWNS, WHOSE SOLUTION INVOLVES AN ARBITRARY PARAMETER*

BY

I. M. SHEFFER

Introduction

This paper gives a discussion of a system of infinitely many linear equations in infinitely many unknowns. The character of the system and of the results differs widely from that of other papers on this subject. It appears that most of the work on such systems is restricted to solutions in Hilbert space: i.e., where $\sum_{i=1}^{\infty} x_i^2$ converges, (x_i) being a solution. The paper of Walsh† may be cited as an exception. He obtains a unique solution, however.

The present work may be described as follows:

In §I we start with a system of linear equations which reduces to a system of difference equations yielding an infinity of solutions. These solutions are shown to be solutions of the original system, after suitable restrictions are laid on the coefficients of the system. We thus get a one-parameter family of solutions. Then we turn, in §II, to a more general system of equations, whose solution can be effected by a method of approximations based on the results of §I. There is shown to exist at least a one-parameter family of solutions (x_i) satisfying the inequality $|x_i| \leqq AP^i$, $P<1$.

In the third section we show that *all* solutions (x_i) which satisfy an inequality $|x_i| \leqq AP^i, P<1$, can be obtained by the method of §II. In §IV there is discussed the method of solution by assuming a solution in the form of a power series, then formally equating coefficients, and finally justifying the formal processes. This method appears to be less effective than the one used in §II.

The writer takes occasion here to express his appreciation for the many suggestions made to him by Professor G. C. Evans.

---

* Presented to the Society, May 3, 1924; received by the editors in October, 1925.

† J. L. Walsh, American Journal of Mathematics, vol. 42 (1920), pp. 91–96.

Some time after submitting this paper to these Transactions the writer's attention was called to the following additional literature:

H. von Koch, *5th International Congress of Mathematicians* (Cambridge, 1912), vol. 1, pp. 352–365.

O. Perron, Mathematische Annalen, vol. 84 (1921), pp. 1–15.

A general reference is to F. Riesz *Les Systèmes d'Equations Linéaires à une Infinité d'Inconnues.*

I. Linear systems reducible to systems of difference equations

Consider the infinite system of linear equations

$$x_i+\lambda \sum_{j=i+1}^{\infty} x_j=c_i \qquad (i=1,2,3,\cdots), \tag{I}$$

where $|c_i|\leqq MP^i$, $P<1$.

Theorem 1. *The system* (I) *has an infinity of solutions for each* λ *in* $|\lambda-1|>1$, *and a single solution for each* λ *in* $|\lambda-1|\leqq 1$.

Proof. On subtracting successive equations we have

$$\begin{aligned} &x_1+(\lambda-1)x_2=c_1-c_2\,, \\ &\cdots\cdots\cdots\cdots \\ &x_n+(\lambda-1)x_{n+1}=c_n-c_{n+1}\,, \\ &\cdots\cdots\cdots\cdots\,. \end{aligned} \tag{II}$$

Every solution of (I) satisfies (II). It remains to be seen when a solution of (II) is also a solution of (I).

Set $\lambda-1=\nu$, and suppose $\nu\neq 0$. Choose $x_1=x_1^{(0)}$ arbitrarily. Then $x_2, x_3, \cdots$ of (II) are uniquely determined. So there are infinitely many solutions of (II) when $\nu\neq 0$. They are given by

$$\begin{aligned} x_1&=x_1^{(0)}\,, \\ x_2&=\frac{-x_1^{(0)}}{\nu}+\frac{c_1-c_2}{\nu}\,, \\ x_3&=\frac{x_1^{(0)}}{\nu^2}-\frac{c_1-c_2}{\nu^2}+\frac{c_2-c_3}{\nu}\,, \\ &\cdots\cdots\cdots\cdots \\ x_n&=(-1)^{n-1}\frac{x_1^{(0)}}{\nu^{n-1}}+(-1)^{n-2}\frac{c_1-c_2}{\nu^{n-1}}+(-1)^{n-3}\frac{c_2-c_3}{\nu^{n-2}}+\cdots+\frac{c_{n-1}-c_n}{\nu}, \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned} \tag{1}$$

$$|c_i|\leqq MP^i\,,\ P<1\,.$$

Therefore

$$\begin{aligned} |x_n|&\leqq\frac{|x_1^{(0)}|}{|\nu|^{n-1}}+\frac{MP(1+P)}{|\nu|^{n-1}}+\frac{MP^2(1+P)}{|\nu|^{n-2}}+\cdots+\frac{MP^{n-1}(1+P)}{|\nu|} \\ &=\frac{|x_1^{(0)}|}{|\nu|^{n-1}}+M(1+P)P^n\left[\frac{1}{|P\nu|^{n-1}}+\frac{1}{|P\nu|^{n-2}}+\cdots+\frac{1}{|P\nu|}\right]. \end{aligned}$$

Assume temporarily that $P|\nu|>1$. (Observe that then $|\nu|>1$.) Then

$$|x_n|\leqq\frac{|x_1^{(0)}|}{|\nu|^{n-1}}+\frac{M(1+P)}{|P\nu|-1}P^n\left[1-\frac{1}{|P\nu|^{n-1}}\right]. \tag{2}$$

The series

$$S_i=x_i+\lambda\sum_{j=i+1}^{\infty}x_j$$

converges absolutely. For

$$|x_i|+|\lambda|\sum_{j=i+1}^{\infty}|x_j|\leqq|x_i|+|\lambda|\left\{\frac{|x_1^{(0)}|}{|\nu|^i}\cdot\frac{1}{1-\frac{1}{|\nu|}}+\frac{M(1+P)}{P|\nu|-1}\cdot\frac{P^{i+1}}{1-P}\right.$$
$$\left.-\frac{MP(1+P)}{(P|\nu|-1)|\nu|^i}\cdot\frac{1}{1-\frac{1}{|\nu|}}\right\}.$$

Let

$$S_{i,n}=x_i+\lambda(x_{i+1}+x_{i+2}+\cdots+x_n).$$

On adding the ith to nth equations of (II) we have

$$S_{i,n}+(\lambda-1)x_{n+1}=c_i-c_{n+1}.$$

Now by hypothesis, $\lim_{n\to\infty}c_n=0$. And from (2) $\lim_{n\to\infty}x_n=0$. Hence $S_i=c_i$. Therefore (x_i) is a solution of (I). We have thus shown that there are an infinity of solutions of (I) for each λ such that $P|\nu|>1$. This conclusion holds for every λ in $|\nu|>1$. For let ν_0 be any value of ν such that $|\nu_0|>1$. We can always choose P' satisfying $1>P'\geqq P$ and such that $P'|\nu_0|>1$. And obviously $|c_i|\leqq MP'^i$. We can therefore replace P by P', and so obtain an infinity of solutions for $\nu=\nu_0$. This proves the first part of the theorem.

Note: In the inequality for $|x_n|$, we must replace P by P' whenever we use P'.

Now assume a solution $x_i(\lambda)$ of (I) in the form of a power series:

$$x_i=A_{0i}+A_{1i}(\lambda-1)+\cdots+A_{ni}(\lambda-1)^n+\cdots. \tag{3}$$

Substituting into (I) formally and equating coefficients, we obtain the following equations to determine A_{ij}:

$$\begin{aligned}&\sum_{j=i}^{\infty}A_{0j}=c_i,\\ &\sum_{j=i}^{\infty}A_{nj}=-\sum_{j=i+1}^{\infty}A_{n-1,j}=c_i^{(n)}.\end{aligned} \tag{4}$$

$A_{0i}=c_i-c_{i+1}$ is a solution of the first system, as is evident. Therefore

$$\overset{(1)}{c_i}=-\sum_{j=i+1}^{\infty}A_{0j}=-\sum_{j=i+1}^{\infty}(c_j-c_{j+1})=-c_{i+1}\ .$$

Hence a solution of

$$\sum_{j=i}^{\infty}A_{1j}=-\sum_{j=i+1}^{\infty}A_{0j}$$

is $A_{1i}=-(c_{i+1}-c_{i+2})$. Therefore

$$\overset{(2)}{c_i}=-\sum_{j=i+1}^{\infty}A_{1j}=c_{i+2}\ .$$

Therefore $A_{2i}=c_{i+2}-c_{i+3}$; and so on. In general, a solution of

$$\sum_{j=i}^{\infty}A_{nj}=c_i^{(n)}$$

is

$$A_{ni}=(-1)^n(c_{i+n}-c_{i+n+1})\ .$$

Therefore

$$\begin{aligned}x_i=(c_i-c_{i+1})-(\lambda-1)(c_{i+1}-c_{i+2})+\cdots\\ +(-1)^n(\lambda-1)^n(c_{i+n}-c_{i+n+1})+\cdots\ ;\end{aligned}$$

that is

$$\begin{aligned}x_i=&\left[c_i-c_{i+1}(\lambda-1)+\cdots+(-1)^n(\lambda-1)^n c_{i+n}+\cdots\right]\\ &-\left[c_{i+1}-c_{i+2}(\lambda-1)+\cdots+(-1)^n(\lambda-1)^n c_{i+n+1}+\cdots\right]\end{aligned}$$

whenever the two series converge. Since $|c_i|\leqq MP^i, P<1$, each series does converge for $|\lambda-1|<1/P$; and

$$|x_i|\leqq\frac{MP^i}{1-P|\nu|}+\frac{MP^{i+1}}{1-P|\nu|}=\frac{M(1+P)}{1-P|\nu|}P^i\ .$$

Therefore

$$|x_i|+|\lambda|\sum_{j=i+1}^{\infty}|x_j|\leqq|x_i|+\frac{M(1+P)|\lambda|}{1-P|\nu|}\cdot\frac{P}{1-P}\cdot P^i\ .$$

Therefore $x_i+\lambda\sum_{j=i+1}^{\infty}x_j$ converges uniformly in every circle of radius $<1/P$. It is therefore legitimate to sum by columns. Consequently

$$x_i+\lambda\sum_{j=i+1}^{\infty}x_j=c_i\ .$$

That is, the formal solution is a true solution. Denote it by $\bar{x}_i$. Set $x_i=y_i+\bar{x}_i$. Then

(5) $$y_i+\lambda\sum_{j=i+1}^{\infty} y_j=0\ .$$

If $\lambda=1$ we see that the unique solution is $y_i=0$. Suppose $\lambda\neq 1$. Then

(6) $$y_i+(\lambda-1)y_{i+1}=0\ , \quad \text{or} \quad y_i=\frac{(-1)^{i-1}}{(\lambda-1)^{i-1}}y_1\ .$$

Add the ith to nth equations in (6):

$$y_i+\lambda(y_{i+1}+\cdots+y_n)=-(\lambda-1)y_{n+1}\ .$$

Hence a solution $y_i\neq 0$ of (6) will be a solution of (5) if and only if $y_n\to 0$ as $n\to\infty$, i.e., if and only if $|\lambda-1|>1$. Hence for $|\lambda-1|\leqq 1$ the only solution of (5) is $y_i=0$, and therefore $\bar{x}_i$ is unique. This completes the proof.

COROLLARY. *Every solution for* $|\lambda-1|>1$ *is obtained by giving all values to* $x_1^{(0)}$.

Consider now the system

(I) $$x_i+\sum_{j=i+1}^{\infty}\lambda_j x_j=c_i \qquad (i=1, 2, \cdots)\ .$$

On subtracting successive equations we obtain

(II) $$x_{i-1}+(\lambda_i-1)x_i=c_{i-1}-c_i\ .$$

Choose $x_1=x_1^{(0)}$ arbitrarily. Then every solution of (II) is given by

$$x_1=x_1^{(0)}\ ,$$

$$\cdots\cdots$$

(1) $$x_n=\frac{(-1)^{n-1}x_1^{(0)}}{\nu_2\nu_3\cdots\nu_n}+\frac{(-1)^{n-2}(c_1-c_2)}{\nu_2\nu_3\cdots\nu_n}+\cdots+\frac{c_{n-1}-c_n}{\nu_n}\ ,$$

where $\nu_i=\lambda_i-1$. Assume that

$$|c_n|\leqq MP^n\ , \quad P<1\ ;$$

(2) $$|\nu_n|\geqq\alpha>\frac{1}{P}\quad(\text{and therefore }\alpha>1)\ ;$$

$$\sum_{j=1}^{\infty}|\lambda_j|P^j \text{ converges.}$$

Then

$$|x_n|\leqq\frac{|x_1^{(0)}|}{\alpha^{n-1}}+M(1+P)P^n\left[\frac{1-\left(\dfrac{1}{\alpha P}\right)^{n-1}}{\alpha P-1}\right]$$

$$\leqq\frac{|x_1^{(0)}|}{\alpha^{n-1}}+\frac{M(1+P)}{\alpha P-1}P^n\ .$$

Therefore

$$\lim_{n\to\infty} x_n = 0 .$$

Also

$$|x_i| + \sum_{j=i+1}^{\infty} |\lambda_j x_j| \leqq |x_i| + \sum_{j=i+1}^{\infty} |\lambda_j| \left\{ \frac{|x_1^{(0)}|}{\alpha^{j-1}} + \frac{M(1+P)}{\alpha P - 1} P^j \right\} ,$$

which converges, by (2). Add the ith to nth equations of (II):

$$x_i + \lambda_{i+1} x_{i+1} + \cdots + \lambda_n x_n = c_i - c_n + x_n .$$

Therefore

$$x_i + \sum_{j=i+1}^{\infty} \lambda_j x_j = c_i + \lim_{n\to\infty} (x_n - c_n) = c_i .$$

Therefore (x_i) is a solution of (I).

Since $x_1^{(0)}$ is arbitrary, there are an infinity of solutions. We have thus

THEOREM 2. *If, in the system* (I) $x_i + \sum_{j=i+1}^{\infty} \lambda_j x_j = c_i$ $(i = 1, 2, 3, \cdots)$, *we have the inequalities* $|c_n| \leqq MP^n$, $P < 1$, *then for each set* (λ_i) *such that* $|\lambda_i - 1| \geqq \alpha > 1/P$ *and* $\sum_{j=1}^{\infty} |\lambda_j| P^j$ *converges, there are an infinity of solutions* (x_i).

COROLLARY. *Every such solution of* (I) *is obtained by giving to* $x_1^{(0)}$ *all values.*

Remarks: (1) Observe that for any such solution (x_i) we have $|x_n| \leqq [|x_1^{(0)}|/P + M(1+P)/(P\alpha - 1)] P^n$; (2) $\sum_{j=1}^{\infty} |\lambda_j| P^j$ converges if $\lim_{j=\infty} |\lambda_{j+1}/\lambda_j| < 1/P$, and therefore λ_j need not be bounded as $j \to \infty$.

II. METHOD OF APPROXIMATIONS APPLIED TO A MORE GENERAL SYSTEM

We now consider the more general system

$$\text{(I)} \qquad x_i + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j = c_i \qquad (i = 1, 2, \cdots) .$$

Assume that

$$|c_i| \leqq MP^i , \quad P < 1 ,$$

$$|\lambda_i - 1| \geqq \alpha > \frac{1}{P} , \quad |b_{ij}| \leqq N ,$$

$$\sum_{j=1}^{\infty} |\lambda_j| P^j \text{ converges.}$$

Let $x_i^{(0,0)}$ be any solution of $x_i + \sum_{j=i+1}^{\infty} \lambda_j x_j = c_i$, and set $x_i^{(1)} = x_i - x_i^{(0,0)}$. By Theorem 2, a solution $x_i^{(0,0)}$ exists. Then, formally,

$$x_i^{(1)} = \{c_i - \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij})x_j\} - \{c_i - \sum_{j=i+1}^{\infty} \lambda_j x_j{}^{(0,0)}\}$$

$$= - \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij})(x_j - x_j{}^{(0,0)}) - \sum_{j=i+1}^{\infty} b_{ij} x_j{}^{(0,0)} ,$$

i.e.,

$$x_i^{(1)} + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j^{(1)} = c_i^{(1)} ,$$

where

$$c_i^{(1)} = - \sum_{j=i+1}^{\infty} b_{ij} x_j^{(0,0)} .$$

Let $x_i^{(0,1)}$ be a solution of

$$x_i + \sum_{j=i+1}^{\infty} \lambda_j x_j = c_i^{(1)} .$$

It will be shown that $c_i^{(1)}$ satisfies the inequality required for Theorem 2, and consequently a solution exists. This remark applies also to the later $c_i^{(n)}$'s: $c_i^{(2)}, c_i^{(3)}, \cdots$. Set

$$x_i^{(2)} = x_i^{(1)} - x_i^{(0,1)} = - \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij})(x_j^{(1)} - x_j^{(0,1)}) - \sum_{j=i+1}^{\infty} b_{ij} x_j^{(0,1)} .$$

Then

$$x_i^{(2)} + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j^{(2)} = c_i^{(2)} ,$$

where

$$c_i^{(2)} = - \sum_{j=i+1}^{\infty} b_{ij} x_j^{(0,1)} ;$$

and so on. In general, let $x_i^{(0,n-1)}$ be a solution of

$$x_i + \sum_{j=i+1}^{\infty} \lambda_j x_j = c_i^{(n-1)} ,$$

and set

$$x_i^{(n)} = x_i^{(n-1)} - x_i^{(0,n-1)} .$$

Then

$$x_i^{(n)} + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j^{(n)} = c_i^{(n)} ,$$

where

$$c_i^{(n)} = - \sum_{j=i+1}^{\infty} b_{ij} x_j^{(0,n-1)} .$$

We need to consider convergence. In solving the system $x_i+\sum_{j=i+1}^{\infty}\lambda_j x_j=c_i$, we obtained the inequality

$$|x_n|\leqq\left[\frac{|x_1^{(0)}|}{P}+\frac{M(1+P)}{P\alpha-1}\right]P^n .$$

In the present case this is

$$|x_n^{(0,0)}|\leqq\left[\frac{|x_1^{(0,0)}|}{P}+\frac{M(1+P)}{P\alpha-1}\right]P^n .$$

Therefore

$$|c_i^{(1)}|\leqq\sum_{j=i+1}^{\infty}|b_{ij}|\,|x_j^{(0,0)}|\leqq N\left[\frac{|x_1^{(0,0)}|}{P}+\frac{M(1+P)}{P\alpha-1}\right]\sum_{j=i+1}^{\infty}P^j$$

$$=N\left[\frac{|x_1^{(0,0)}|}{P}+\frac{M(1+P)}{P\alpha-1}\right]\frac{P}{1-P}\cdot P^i .$$

Let

$$Q=\frac{N}{1-P}, \qquad R=\frac{PN(1+P)}{(1-P)(P\alpha-1)} ,$$

$$M^{(1)}=Q\,|x_1^{(0,0)}|+RM .$$

Then

$$|c_i^{(1)}|\leqq M^{(1)}P^i .$$

Assume α chosen so that $R<1$. This is equivalent to assuming that

$$\alpha>\frac{NP^2+(N-1)P+1}{P(1-P)} .$$

(Observe that this inequality implies $P\alpha>1$.) In the solution $x_i^{(0,1)}$ of $x_i+\sum_{j=i+1}^{\infty}\lambda_j x_j=c_i^{(1)}$, $x_1^{(0,1)}$ can be taken arbitrarily (see Theorem 2). Therefore

$$|x_n^{(0,1)}|\leqq\left[\frac{|x_1^{(0,1)}|}{P}+\frac{M^{(1)}(1+P)}{P\alpha-1}\right]P^n .$$

Therefore

$$|c_i^{(2)}|\leqq\sum_{j=i+1}^{\infty}|b_{ij}|\,|x_j^{(0,1)}|\leqq N\left[\frac{|x_1^{(0,1)}|}{P}+\frac{M^{(1)}(1+P)}{P\alpha-1}\right]\frac{P}{1-P}P^i$$

$$=[Q\,|x_1^{(0,1)}|+RM^{(1)}]P^i$$

$$=M^{(2)}P^i\text{ , say;}$$

and so on. It is clear that in general

$$|x_n^{(0,r)}| \leqq \left[\frac{|x_1^{(0,r)}|}{P} + \frac{M^{(r)}(1+P)}{P\alpha-1}\right]P^n ,$$

and

$$|c_i^{(r+1)}| \leqq \sum_{j=i+1}^{\infty} |b_{ij}|\,|x_j^{(0,r)}|$$

$$\leqq [Q\,|x_1^{(0,r)}|+RM^{(r)}]P^i = M^{(r+1)}P^i ,$$

where $x_1^{(0,r)}$ is chosen arbitrarily, and $M^{(r)}$ is defined by the relation

$$M^{(r+1)} = Q\,|x_1^{(0,r)}|+RM^{(r)} ,$$

$$M^{(0)} = M .$$

From this recurrent relation for $M^{(r)}$ we have

$$M^{(r)} = Q\,|x_1^{(0,r-1)}|+QR\,|x_1^{(0,r-2)}|+QR^2\,|x_1^{(0,r-3)}|$$

$$+ \cdots + QR^{r-1}|x_1^{(0,0)}| + R^rM.$$

The choice of $x_1^{(0,0)}$, $x_1^{(0,1)}$, $\cdots$, $x_1^{(0,r)}$, $\cdots$ is arbitrary. Now *assume* them so chosen that

$$|x_1^{(0,r)}| \leqq \beta T^r , \qquad R<T<1 .$$

Then

$$M^{(r)} \leqq \beta Q T^{r-1}\left[1+\frac{R}{T}+\frac{R^2}{T^2}+\cdots+\frac{R^{r-1}}{T^{r-1}}\right]+MR^r$$

$$= \beta Q T^{r-1}\left[\frac{1-\left(\dfrac{R}{T}\right)^r}{1-\dfrac{R}{T}}\right]+MR^r$$

$$\leqq \left[\frac{\beta Q}{1-\dfrac{R}{T}}+MR\right]T^{r-1} = ET^{r-1} , \quad \text{say} .$$

Therefore $|c|_i^{(r+1)} \leqq ET^rP^i$, and

$$|x_n^{(0,r)}| \leqq \left[\frac{\beta T^r}{P}+\frac{(1+P)}{P\alpha-1}ET^{r-1}\right]P^n$$

$$= \left[\beta\frac{T}{P}+\frac{(1+P)E}{P\alpha-1}\right]T^{r-1}P^n = HT^{r-1}P^n , \text{ say.}$$

Observe that $\lim_{n\to\infty} c_i^{(n)}=0$. Let $X_i=\sum_{r=0}^{\infty} x_i^{(0,r)}$. The series converges absolutely, since

$$\sum_{r=0}^{\infty} |x_i^{(0,r)}| \leqq HP^i \sum_{r=0}^{\infty} T^{r-1}\,, \qquad T<1\,.$$

Define $Y_i^{(n)}$ by

$$X_i = x_i^{(0,0)} + x_i^{(0,1)} + \cdots + x_i^{(0,n-1)} + Y_i^{(n)}\,.$$

Then

$$|Y_i^{(n)}| \leqq \sum_{r=n}^{\infty} |x_i^{(0,r)}| \leqq HP^i \sum_{r=n}^{\infty} T^{r-1}\,.$$

Therefore

$$\lim_{n\to\infty} Y_i^{(n)} = 0\,.$$

We have

$$\sum_{j=i+1}^{\infty} |\lambda_j+b_{ij}|\,|x_j^{(0,r)}| \leqq NHT^{r-1} \sum_{j=i+1}^{\infty} P^j + HT^{r-1} \sum_{j=i+1}^{\infty} |\lambda_j| P^j\,.$$

Both series on the right converge, and therefore $\sum_{j=i+1}^{\infty}(\lambda_j+b_{ij})x_j^{(0,r)}$ converges absolutely for every r. Also

$$\sum_{j=i+1}^{\infty} |\lambda_j+b_{ij}|\,|Y_j^{(n)}| \leqq NH\left(\sum_{r=n}^{\infty} T^{r-1}\right)\sum_{j=i+1}^{\infty} P^j$$
$$+H\left(\sum_{r=n}^{\infty} T^{r-1}\right)\sum_{j=i+1}^{\infty} |\lambda_j| P^j\,,$$

and both series on the right converge. Therefore $\sum_{j=i+1}^{\infty}(\lambda_j+b_{ij})Y_j^{(n)}$ converges absolutely. Moreover,

$$\lim_{n\to\infty} \sum_{j=i+1}^{\infty} (\lambda_j+b_{ij}) Y_j^{(n)} = 0\,.$$

Therefore

$$\begin{aligned} X_i &+ \sum_{j=i+1}^{\infty} (\lambda_j+b_{ij}) X_j \\ &= x_i^{(0,0)} + \sum_{j=i+1}^{\infty} (\lambda_j+b_{ij}) x_j^{(0,0)} \\ &\quad + x_i^{(0,1)} + \sum_{j=i+1}^{\infty} (\lambda_j+b_{ij}) x_j^{(0,1)} \\ &\quad + \cdots + x_i^{(0,n-1)} + \sum_{j=i+1}^{\infty} (\lambda_j+b_{ij}) x_j^{(0,n-1)} \end{aligned}$$

$$+Y_i^{(n)}+\sum_{j=i+1}^{\infty}(\lambda_j+b_{ij})Y_j^{(n)}$$

$$=(c_i-c_i^{(1)})+(c_i^{(1)}-c_i^{(2)})+\cdots+(c_i^{(n-1)}-c_i^{(n)})$$

$$+Y_i^{(n)}+\sum_{j=i+1}^{\infty}(\lambda_j+b_{ij})Y_j^{(n)}$$

$$=c_i+\Big[-c_i^{(n)}+Y_i^{(n)}+\sum_{j=i+1}^{\infty}(\lambda_j+b_{ij})Y_j^{(n)}\Big].$$

As $n\to\infty$, the bracket approaches 0.

Now the left hand member is independent of n. Therefore

$$X_i+\sum_{j=i+1}^{\infty}(\lambda_j+b_{ij})X_j=c_i\ ;$$

i.e., (X_i) is a solution of (I).

Except for the condition $|x_1^{(0,r)}|\leqq\beta T^r$, $R<T<1$, the $x_1^{(0,r)}$'s are arbitrary. Therefore X_1 is arbitrary. Consequently there are an infinity of solutions. We thus have

THEOREM 3. *If in the system*

$$\text{(I)}\qquad x_i+\sum_{j=i+1}^{\infty}(\lambda_j+b_{ij})x_j=c_i\qquad (i=1,2,3,\cdots),$$

the conditions $|c_n|\leqq MP^n$, $P<1$, *and* $|b_{ij}|\leqq N$ *hold, then to every set* (λ_i) *for which* $|\lambda_i-1|\geqq\alpha>(NP^2+(N-1)P+1)/P(1-P)$ *and* $\sum_{j=1}^{\infty}|\lambda_j|P^j$ *converges, there exist an infinity of solutions of* (I).

COROLLARY. *From the inequality*

$$|X_i|\leqq HP^i\sum_{r=0}^{\infty}T^{r-1}$$

we have the following result: An infinity of the solutions (for a given set λ_i*) satisfy the inequality* $|x_i|\leqq AP^i$, *where* A *is a constant (depending on* α*).*

COROLLARY. *If* $\lambda_j\equiv\lambda$, *then the system* $x_i+\sum_{j=i+1}^{\infty}(\lambda+b_{ij})x_j=c_i$ *has an infinity of solutions for each* λ *in* $|\lambda-1|>(NP^2+(N-1)P+1)/P(1-P)$, *if* $|c_i|\leqq MP^i$, $P<1$, $|b_{ij}|\leqq N$, *and an infinity of these solutions satisfy* $|x_i|\leqq AP^i$, *A depending on* λ.

COROLLARY. *Consider the system* $y_i+\sum_{j=i+1}^{\infty}(\lambda_j+b_{ij})y_j=c_i$, *where* $|b_{ij}|\leqq N$; *P exists,* $0<P<1$, *such that* $\sum_{j=1}^{\infty}|\lambda_j|P^j$ *converges;* $|\lambda_i-1|\geqq\alpha>(NP^2+(N-1)P+1)/P(1-P)$. *If this system has one solution, then it has an infinite number.*

For let $\bar{y}_i$ be a solution, and set $y_i = x_i + \bar{y}_i$. Then $x_i + \sum_{j=i+1}^{\infty}(\lambda_j + b_{ij}) = 0$, which is a system satisfying the conditions of Theorem 3. Therefore an infinity of solutions x_i, and consequently y_i, exist.

III. Solutions satisfying the inequality $|x_i| \leqq AP^i$*

In obtaining Theorem 3, we assumed that

$$|x_1^{(0,r)}| \leqq \beta T^r , \qquad R < T < 1 .$$

Let us now make a more particular choice; take

$$x_1^{(0,r)} = 0 \qquad (r = 0, 1, 2, \cdots).$$

Then

$$M^{(r)} = MR^r ,$$

$$|c_i^{(r)}| \leqq MR^r P^i ,$$

$$|x_n^{(0,r)}| \leqq \frac{M(1+P)}{P\alpha - 1} R^r P^n .$$

Therefore

$$|X_i| = |\sum_{r=0}^{\infty} x_i^{(0,r)}| \leqq \frac{M(1+P)}{P\alpha-1}\left(\sum_{r=0}^{\infty} R^r\right)P^i = BMP^i ,$$

where

$$B = \frac{1+P}{(P\alpha-1)(1-R)} .$$

We shall use these inequalities.

Consider again System (I) of Theorem 3. We found that for a given permissible set (λ_i) an infinity of solutions exist satisfying

$$|X_i| \leqq AP^i .$$

We shall now show that *every* solution satisfying

$$|X_i| \leqq AP^i ,$$

A a constant, is obtained by the method given in Theorem 3. We shall however require a modification in the inequality for α.

Let $X_i^{(0,0)}$ be a solution for which $|X_i^{(0,0)}| \leqq AP^i$. Then

$$X_i^{(0,0)} + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) X_j^{(0,0)} = c_i .$$

* The method used in this section is due to G. C. Evans, these Transactions, vol. 12 (1911), pp. 452–457.

Consider the system

$$x_i + \sum_{j=i+1}^{\infty} \lambda_j x_j = c_i \,.$$

Define

$$x_i^{(1)} = x_i - X_i^{(0,0)} \,.$$

Then $x_i^{(1)} + \sum_{j=i+1}^{\infty} \lambda_j x_j^{(1)} = c_i^{(1)}$, where $c^{(1)} = \sum_{j=i+1}^{\infty} b_{ij} X_j^{(0,0)}$. The series for $c_i^{(1)}$ obviously converges absolutely, and

$$|c_i^{(1)}| \leqq A \frac{NP}{1-P} \cdot P^i \,, \quad = M^{(1)} P^i \,, \text{ say.}$$

Now let $X_i^{(0,1)}$ be a solution of

$$x_i + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j = c_i^{(1)}$$

corresponding to the choice $x_1^{(0,r)} = 0$, $r = 0, 1, 2, \cdots$ in Theorem 3. Then $|X_i^{(0,1)}| \leqq BM^{(1)}P^i$. Define $x_i^{(2)} = x_i^{(1)} - X_i^{(0,1)}$. Then $x_i^{(2)} + \sum_{j=i+1}^{\infty} \lambda_j x_j^{(2)} = c_i^{(2)}$, where $c_i^{(2)} = \sum_{j=i+1}^{\infty} b_{ij} X_j^{(0,1)}$. Then

$$|c_i^{(2)}| \leqq \frac{NBM^{(1)}P}{1-P} \cdot P^i = BA \left(\frac{NP}{1-P} \right)^2 P^i = M^{(2)} P^i \,.$$

In general,

$$x_i^{(n)} + \sum_{j=i+1}^{\infty} \lambda_j x_j^{(n)} = c_i^{(n)} \,,$$

where

$$c_i^{(n)} = \sum_{j=i+1}^{\infty} b_{ij} X_j^{(0,n-1)} \,, \qquad x_i^{(n)} = x_i^{(n-1)} - X_i^{(0,n-1)} \,,$$

and $X_i^{(0,n-1)}$ is a solution of

$$x_i + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j = c_i^{(n-1)} \,,$$

corresponding to the choice $x_1^{(0,r)} = 0$, $r = 0, 1, 2, \cdots$; and

$$|c_i^{(r)}| \leqq M^{(r)} P^i \,,$$

where

$$M^{(r)} = \frac{NBP}{1-P} M^{(r-1)} \,, \qquad M^{(1)} = \frac{NAP}{1-P} \,.$$

Hence

$$M^{(r)} = \frac{A}{B} \left(\frac{BNP}{1-P} \right)^r = \frac{A}{B} W^r \,,$$

where

$$W=\frac{BNP}{1-P}.$$

Also,

$$|X_i^{(0,r)}|\leqq BM^{(r)}P^i=AW^rP^i.$$

Assume $W<1$. This is equivalent to assuming that $R<\frac{1}{2}$; i.e., that

$$\alpha>\frac{2NP^2+(2N-1)P+1}{P(1-P)}.$$

Observe that this inequality implies the previous one,

$$\alpha>\frac{NP^2+(N-1)P+1}{P(1-P)}.$$

Since $W<1$, $\lim_{n\to\infty}c_i^{(n)}=0$.

Define $\bar{x}_i^{(0,0)}$ by the series

$$\bar{x}_i^{(0,0)}=\sum_{r=0}^{\infty}X_i^{(0,r)}.$$

The series converges absolutely, and

$$|\bar{x}_i^{(0,0)}|\leqq AP^i\sum_{r=0}^{\infty}W^r=\frac{A}{1-W}P^i.$$

Let

$$y_i^{(n)}=X_i^{(0,n+1)}+X_i^{(0,n+2)}+\cdots.$$

Then

$$|y_i^{(n)}|\leqq AP^i\sum_{r=n+1}^{\infty}W^r;$$

therefore $\lim_{n\to\infty}y_i^{(n)}=0$. Again, from the inequality on $|X_j^{(0,r)}|$ we have

$$\sum_{j=i+1}^{\infty}|\lambda_j|\,|X_j^{(0,r)}|\leqq AW^r\sum_{j=i+1}^{\infty}|\lambda_j|P^j,$$

which converges; hence

$$\sum_{j=i+1}^{\infty}|\lambda_j|\,|y_j^{(n)}|\leqq A\left(\sum_{r=n+1}^{\infty}W^r\right)\sum_{j=i+1}^{\infty}|\lambda_j|P^j.$$

Therefore $\lim_{n\to\infty}\sum_{j=i+1}^{\infty}\lambda_j y_j^{(n)}=0$. Then

$$\bar{x}_i^{(0,0)} + \sum_{j=i+1}^{\infty} \lambda_j \bar{x}_j^{(0,0)} = X_i^{(0,0)} + \sum_{j=i+1}^{\infty} \lambda_j X_j^{(0,0)} + \cdots + X_i^{(0,n)}$$

$$+ \sum_{j=i+1}^{\infty} \lambda_j X_j^{(0,n)} + y_i^{(n)} + \sum_{j=i+1}^{\infty} \lambda_j y_j^{(n)}$$

$$= (c_i - c_i^{(1)}) + (c_i^{(1)} - c_i^{(2)}) + \cdots + (c_i^{(n)} - c_i^{(n+1)})$$

$$+ y_i^{(n)} + \sum_{j=i+1}^{\infty} \lambda_j y_j^{(n)}$$

$$= c_i + [-c_i^{(n+1)} + y_i^{(n)} + \sum_{j=i+1}^{\infty} \lambda_j y_j^{(n)}] .$$

The bracket approaches zero as $n \to \infty$. Therefore $\bar{x}_i^{(0,0)} + \sum_{j=i+1}^{\infty} \lambda_j \bar{x}_j^{(0,0)} = c_i$, i.e., $(\bar{x}_i^{(0,0)})$ is a solution of $x_i + \sum_{j=i+1}^{\infty} \lambda_j x_j = c_i$.

Now we shall start with the solution $\bar{x}_i^{(0,0)}$ and show that one of the solutions X_i of

$$x_i + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j = c_i$$

resulting from the method of Theorem 3 is precisely $X_i^{(0,0)}$.

Define $\bar{x}_i^{(0,1)}, \bar{x}_i^{(0,2)}, \cdots$ as solutions of the systems

$$\bar{x}_i^{(0,1)} + \sum_{j=i+1}^{\infty} \lambda_j \bar{x}_j^{(0,1)} = - \sum_{j=i+1}^{\infty} b_{ij} \bar{x}_j^{(0,0)} = e_i^{(1)} ,$$

$$\bar{x}_i^{(0,2)} + \sum_{j=i+1}^{\infty} \lambda_j \bar{x}_j^{(0,2)} = - \sum_{j=i+1}^{\infty} b_{ij} \bar{x}_j^{(0,1)} = e_i^{(2)} ,$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

respectively. There are of course an infinite number of solutions of each of these systems, by Theorem 2. The precise solutions which we choose will be made evident later.

$X_i^{(0,0)}$ is a solution of $x_i + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j = c_i$, and

$$|X_i^{(0,0)}| \leqq A P^i .$$

Let $X_i^{(1,0)} = - \sum_{r=1}^{\infty} X_i^{(0,r)}$; then $|X_i^{(1,0)}| \leqq A(W/(1-W))P^i$.

Then $X_i^{(1,0)}$ satisfies $x_i + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j = e_i^{(1)}$. For $X_i^{(1,0)} = X_i^{(0,0)} - \bar{x}_i^{(0,0)}$, and

$$X_i^{(0,0)} + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) X_j^{(0,0)} = c_i , \qquad \bar{x}_i^{(0,0)} + \sum_{j=i+1}^{\infty} \lambda_j \bar{x}_j^{(0,0)} = c_i .$$

We have

$$|e_i^{(1)}| \leqq \frac{NAP}{(1-W)(1-P)} P^i .$$

Now define $X_i^{(1,1)}, X_i^{(1,2)}, \cdots, X_i^{(1,n)}, \cdots$ with respect to $X_i^{(1,0)}$ in the same way as $X_i^{(0,1)}, \cdots, X_i^{(0,n)}, \cdots$ were defined with respect to $X_i^{(0,0)}$. Then

$$|X_i^{(1,r)}| \leqq A\left(\frac{W}{1-W}\right) W^r P^i .$$

Choose $\bar{x}_i^{(0,1)}$ as that solution of

$$x_i + \sum_{j=i+1}^{\infty} \lambda_j x_j = e_i^{(1)}$$

corresponding to the set $X_i^{(1,r)}$:

$$\bar{x}_i^{(0,1)} = \sum_{r=0}^{\infty} X_i^{(1,r)} .$$

Now define $X_i^{(2,0)} = -\sum_{r=1}^{\infty} X_i^{(1,r)}, = X_i^{(1,0)} - \bar{x}_i^{(0,1)}$. Then $X_i^{(2,0)}$ is a solution of

$$x_i + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j = e_i^{(2)} ,$$

and

$$|e_i^{(2)}| \leqq \frac{NAP}{(1-W)(1-P)} \left(\frac{W}{1-W}\right) P^i ; \qquad |X_i^{(2,0)}| \leqq A\left(\frac{W}{1-W}\right)^2 P^i .$$

Define the corresponding set $X_i^{(2,1)}, X_i^{(2,2)}, \cdots$; then

$$|X_i^{(2,r)}| \leqq A\left(\frac{W}{1-W}\right)^2 W^r P^i .$$

Then choose $\bar{x}_i^{(0,2)}$ as that solution of

$$x_i + \sum_{j=i+1}^{\infty} \lambda_j x_j = e_i^{(2)}$$

which is determined by $X_i^{(2,r)}$:

$$\bar{x}_i^{(0,2)} = \sum_{r=0}^{\infty} X_i^{(2,r)} ;$$

and so on. In general, we define

$$X_i^{(k,0)} = -\sum_{r=1}^{\infty} X_i^{(k-1,r)} , \quad = X_i^{(k-1,0)} - \bar{x}_i^{(0,k-1)} .$$

Then $X_i^{(k,0)}$ satisfies $x_i+\sum_{j=i+1}^{\infty}(\lambda_j+b_{ij})x_j=e_i^{(k)}$, where $e_i^{(k)}=-\sum_{j=i+1}^{\infty} b_{ij}\bar{x}_j^{(0,k-1)}$.

By induction we have $|X_i^{(k,0)}|\leqq A(W/(1-W))^kP^i$. We then define $X_i^{(k,1)}$, $X_i^{(k,2)}, \cdots$, and obtain $|X_i^{(k,r)}|\leqq A(W/(1-W))^kW^rP^i$. Also

$$|e_i^{(k)}|\leqq\frac{NAP}{(1-P)(1-W)}\left(\frac{W}{1-W}\right)^{k-1}P^i .$$

Choose $\bar{x}_i^{(0,k)}$ as that solution of $x_i+\sum_{j=i+1}^{\infty}\lambda_jx_j=e_i^{(k)}$ which is given by $\bar{x}_i^{(0,k)}=\sum_{r=0}^{\infty}X_i^{(k,r)}$. Observe that

$$|\bar{x}_1^{(0,k)}|\leqq\sum_{r=0}^{\infty}A\left(\frac{W}{1-W}\right)^kW^rP=\frac{AP}{1-W}\left(\frac{W}{1-W}\right)^k .$$

This is in the form $|\bar{x}_1^{(0,k)}|\leqq\beta T'^k$, $T'<1$, provided $W/(1-W)<1$; i.e., $W<\frac{1}{2}$, or $R<\frac{1}{3}$, which is equivalent to the inequality

$$\alpha>\frac{3NP^2+(3N-1)P+1}{P(1-P)} .$$

We *assume* this condition on α. (It implies the previous condition $R<\frac{1}{2}$.) Consequently, by the method of Theorem 3,

$$X_i=\sum_{r=0}^{\infty}\bar{x}_i^{(0,r)}$$

is a solution of $x_i+\sum_{j=i+1}^{\infty}(\lambda_j+b_{ij})x_j=c_i$.

It remains to identify X_i with $X_i^{(0,0)}$. We see that

$$\begin{aligned}\bar{x}_i^{(0,0)}&+\bar{x}_i^{(0,1)}+\cdots+\bar{x}_i^{(0,n)}\\ &=X_i^{(0,0)}+\{X_i^{(0,1)}+X_i^{(0,2)}+\cdots\}\\ &\quad+X_i^{(1,0)}+\{X_i^{(1,1)}+X_i^{(1,2)}+\cdots\}+\cdots\\ &\quad+X_i^{(n,0)}+\{X_i^{(n,1)}+\cdots\},\end{aligned}$$

where the right member, from the definition of $X_i^{(k,0)}$, $k=1, 2, \cdots, n+1$, reduces to

$$X_i^{(0,0)}-X_i^{(n+1,0)} .$$

Now $\lim_{n\to\infty}X_i^{(n+1,0)}=0$. Therefore

$$\lim_{n\to\infty}[\bar{x}_i^{(0,0)}+\cdots+\bar{x}_i^{(0,n)}]=X_i^{(0,0)} ;$$

that is,

$$X_i = X_i^{(0,0)} \,.$$

We have thus proved

THEOREM 4. *If in the system*

$$\text{(I)} \qquad x_i + \sum_{j=i+1}^{\infty} (\lambda_j + b_{ij}) x_j = c_i \qquad (i = 1, 2, \cdots),$$

the following inequalities hold:

$$|c_i| \leqq MP^i, \qquad P<1, \qquad |b_{ij}| \leqq N;$$

$$|\lambda_i - 1| \geqq \alpha > \frac{3NP^2 + (3N-1)P + 1}{P(1-P)};$$

and $\sum_{j=i+1}^{\infty} |\lambda_j| P^j$ *converges, then every solution* x_i *for which* $|x_i| \leqq AP^i$ *(of which there are an infinite number) is contained in the set obtained by the method of Theorem* 3, *assigning to* $x_1^{(0,0)}, x_1^{(0,1)}, \cdots$ *suitable values in the range* $|x_1^{(0,r)}| \leqq \beta T^r$, $T<1$.

IV. THE POWER SERIES METHOD

In Theorem 1 we made use of a Taylor series in λ, and found a solution by substituting the assumed series for x_i into the system and equating coefficients. This method can be applied to other systems.

For example, consider the system

$$\text{(I)} \qquad x_i + \lambda \sum_{j=i+1}^{\infty} \alpha_{ij} x_j = c_i \,.$$

Assume a solution $x_i = A_{0i} + \cdots + A_{ni}\lambda^n + \cdots$. We find for the A_{ij}'s the relations

$$A_{0i} = c_i \,,$$

$$A_{ni} = -\sum_{j=i+1}^{\infty} \alpha_{ij} A_{n-1,j} \,, \qquad n>0 \,.$$

It is readily shown that if $|c_i| \leqq C$, $\sum_{j=i+1}^{\infty} |\alpha_{ij}| \leqq P$, $i \geqq i_0$, then (I) has a bounded solution (x_i) for every λ in $|\lambda| < 1/P$. But this method does not appear to yield uniqueness properties.

This method also gives the following theorem:

If

$$|c_i| \leqq \alpha\ ; \quad \sum_{j=i+1}^{\infty} |\alpha_{ij}| \leqq S \qquad (i=1, 2, \cdots)\ ;$$

$$\varphi_{ij}(x) \textit{ is analytic}\ , \quad |x| \leqq M\ , \quad M > \alpha$$

$$(i=1, 2, \cdots\ ; \quad j=i+1, i+2, \cdots)\ ;$$

$$|\varphi_{ij}(x)| \leqq N\ , \quad |x| \leqq M\ ;$$

then the system

$$x_i + \lambda \sum_{j=i+1}^{\infty} \alpha_{ij}\varphi_{ij}(x_j) = c_i$$

has a bounded solution for every λ *in* $|\lambda| \leqq (M-\alpha)/5NS$.

It is an interesting fact that if we assume a power series solution in $1/\lambda$, we obtain in many cases an infinite number of solutions. Let us consider again the system of Theorem 3, where however we take $\lambda_i \equiv \lambda$:

$$x_i + \sum_{j=i+1}^{\infty} (\lambda + b_{ij})x_j = c_i \begin{cases} |c_i| \leqq MP^i\ , & P < 1\ , \\ |b_{ij}| \leqq N\ . \end{cases}$$

Assume a solution

$$x_i = A_{0i} + \cdots + \frac{A_{ni}}{(\lambda-1)^n} + \cdots .$$

On substituting in and equating coefficients we get

$$\sum_{j=i+1}^{\infty} A_{0j} = 0\ , \qquad A_{0i} + \sum_{j=i+1}^{\infty} (b_{ij}+1)A_{0j} + \sum_{j=i+1}^{\infty} A_{1j} = c_i\ ,$$

$$A_{ni} + \sum_{j=i+1}^{\infty} (b_{ij}+1)A_{nj} + \sum_{j=i+1}^{\infty} A_{n+1,j} = 0\ , \qquad n > 0\ .$$

A solution of the first system is $A_{0i}=0$, $i>1$. Observe that A_{01} does not enter, and therefore can be taken arbitrarily. And in the general system, A_{n1} does not enter and is therefore arbitrary. This makes x_1 arbitrary, and we therefore obtain an infinity of solutions. Let us determine the A_{ij}'s. We have the relation

$$\sum_{j=i+1}^{\infty} A_{1j} = c_i - A_{0i} = c_i^{(1)}\ .$$

Therefore a solution is $A_{1,i+1} = c_i^{(1)} - c_{i+1}^{(1)}$. Also

$$\sum_{j=i+1}^{\infty} A_{2j} = -A_{1i} - \sum_{j=i+1}^{\infty} (b_{ij}+1) A_{1j} = c_i^{(2)},$$

and a solution is

$$A_{2,i+1} = c_i^{(2)} - c_{i+1}^{(2)}.$$

In general $A_{n,i+1} = c_i^{(n)} - c_{i+1}^{(n)}$, where $c_i^{(n)} = -A_{n-1,i} - \sum_{j=i+1}^{\infty} (b_{ij}+1) A_{n-1,j}$. By definition, $c_i^{(1)} = c_i - A_{0i}$. Choose A_{01} so that $|c_1 - A_{01}| \leqq MP$; otherwise arbitrary. Then $|c_i^{(1)}| \leqq MP^i = M^{(1)}P^i$. Therefore $|A_{1,i+1}| \leqq M^{(1)}(1+P)P^i$. Choose A_{11} arbitrary except for the condition $|A_{11}| \leqq M(1+P)$. Then

$$|A_{1i}| \leqq M(1+P)P^{i-1} = \frac{M(1+P)}{P} P^i.$$

Therefore

$$|c_i^{(2)}| \leqq \frac{M^{(1)}(1+P)}{P}\left[1+\frac{(N+1)P}{1-P}\right]P^i = M^{(2)}P^i.$$

Therefore $|A_{2,i+1}| \leqq M^{(2)}(1+P)P^i$. Choose $|A_{2,1}| \leqq M^{(2)}(1+P)$; then $|A_{2i}| \leqq (M^{(2)}(1+P)/P)P^i$; and so on. We obtain, finally, $|c_i^{(n)}| \leqq M^{(n)}P^i$, where

$$M^{(n)} = RM^{(n-1)} = \cdot \cdot \cdot = MR^{n-1},$$

$$R = \left(1+\frac{P(N+1)}{1-P}\right)\left(\frac{1+P}{P}\right).$$

Therefore

$$|c_i^{(n)}| \leqq MR^{n-1}P^i,$$

$$|A_{ni}| \leqq \frac{M(1+P)}{P} R^{n-1} P^i,$$

on choosing

$$|A_{n1}| \leqq MR^{n-1}(1+P).$$

Therefore

$$|x_i| \leqq P^i \frac{M(1+P)}{RP} \sum_{n=0}^{\infty} \frac{R^n}{|\lambda-1|^n} = T_i,$$

which converges if $|\lambda-1| > R$; i.e., if

$$|\lambda-1| > \frac{(1+P)(NP+1)}{P(1-P)}.$$

Also, $T_i + \sum_{j=i+1}^{\infty} |\lambda+b_{ij}| T_j$ converges. Hence we can sum by columns. We thus have the following theorem:

If $|c_i| \leqq MP^i$, $P<1$, $|b_{ij}| \leqq N$, *and* $|\lambda - 1| > (1+P)(NP+1)/P(1-P)$, *then the system* $x_i + \sum_{j=i+1}^{\infty}(\lambda + b_{ij})x_j = c_i$ *has an infinity of solutions.*

So we obtain some of the results of Theorem 3. But the power series method does not extend to the case where the λ_j's are not all equal; and even in the case just treated, the λ- region for which the proof is valid is not as extended as the region found by the method of Theorem 3.

RICE INSTITUTE,
HOUSTON, TEXAS.

SOLUTION OF CERTAIN FUNCTIONAL EQUATIONS RELATIVE TO A GENERAL LINEAR SET*

BY

MARK H. INGRAHAM

I. Introduction

This paper exhibits the solution of the functional equation

$$(A) \qquad \sum_{j}^{0,n+1} (-1)^{j}_{n+1}C_j\, f(x_0+jx_1)=0$$

for every x_0 and x_1 belonging to a linear set of a certain type. As is well known, relative to the complex, real or rational number system any polynomial of degree n satisfies (A). Equation (A) simply states that the $(n+1)$st difference of the function f vanishes for every set of values of the argument with a constant difference.

Both the equation studied and the method pursued are generalizations of the equation and method of Legendre,† Hamel‡ and Schimmack§ in their studies of the equation

$$f(x+y)=f(x)+f(y)\ .$$

It is obvious that any solution of (1) is also a solution of (A) for the case in which $n=1$.

Hamel shows that if we assume that the system $\mathfrak{R}$ of real numbers can be normally ordered (well ordered) then a set $\mathfrak{R}_0$ of real numbers exists, such that all elements of $\mathfrak{R}_0$ are linearly independent as to finite combinations with rational coefficients and such that every real number is a linear combination of a finite number of elements of $\mathfrak{R}_0$ with rational coefficients.

Consider $\mathfrak{R}_0$ as a range and V the set of all vectors with rational coördinates of which, for any one vector, only a finite number are different from zero. It is obvious that the elements of $\mathfrak{R}$ may be set into one to one correspondence with the elements of V in such a way that the correspondence is preserved under addition and multiplication by rational numbers.

Hamel shows that any solution of (1) is a linear combination of the coördinates of the vectors of V. This paper shows that as a direct generaliza-

* Presented to the Society, April 11, 1925; received by the editors before December, 1925.

† A. M. Legendre (not verified by direct reference) *Eléments de Géométrie*, 4th edition, Paris, 1802, p. 279.

‡ Hamel, Mathematische Annalen, vol. 60 (1905), p. 459.

§ R. Schimmack, *Axiomatische Untersuchungen über die Vektoraddition*, Nova Acta, Halle, 1908.

tion the solution of (A) is a polynomial in the coördinates of the vectors of V.

II. Résumé of certain results in the theory of linear sets*

Consider a number system $\mathfrak{A} \equiv [a]$ which is of such a type† that it permits of unique associative and commutative addition, and of unique associative but not necessarily commutative multiplication and also permits of subtraction and division by all numbers not zero while multiplication is distributive as to addition. Any field is such a system, likewise the totality of all real quaternions forms such a system. Such a system $\mathfrak{A}$ we say is of type A.

Consider $\mathfrak{A}' = [\text{all } a \,.\ni.\, a_1 .).\, aa_1 = a_1 a]$,‡ i. e., the totality of all members of $\mathfrak{A}$ which are commutative as to multiplication with every member of $\mathfrak{A}$. Then $\mathfrak{A}'$ is a field.§

Every field contains a sub-field R isomorphic with the field of all rational numbers or the field of all integers modulo a prime, p. In the first case we will say that R and consequently $\mathfrak{A}$ and $\mathfrak{A}'$ have the modulus ω, otherwise the modulus p. The systems R and $\mathfrak{A}$ form an example of a system $\mathfrak{A}$ of type A and an abstract set of vectors‖ U connected by a unique associative and commutative addition process and connected with $\mathfrak{A}$ by a process of scalar multiplication which is unique, associative and distributive with respect to addition but not necessarily commutative. In this particular case scalar multiplication is commutative.

If we assume that U has a normally ordered base U_1 in terms of the elements of which any element of U is expressible as a finite linear combination with coefficients in R, it can be shown that there exists a base U_0 having the same property and such that its terms are finitely linearly independent as to coefficients in R. In such a case U is isomorphic with the totality $V = [v]$ of vectors on U_0 to R finitely non-zero. Of course, if U is normally ordered U itself is effective as U_1.¶

---

* See my paper on *A general theory of linear sets*, these Transactions, vol. 27 (1925), pp. 163–196. The facts summarized in this section are given in detail with proofs in that paper.

† Ibid., p. 164 ff.

‡ Throughout this paper the logical symbols for implies, such that, etc., are used. Though no system of logical symbolism can yet be called classical, yet several are so similar that any one familiar with one may readily read the others. In all essentials the notation here used is the same as that used by E. H. Moore in his *Introduction to a Form of General Analysis*, Yale University Press, 1910. See p. 150. The definitions are also given in my paper, loc. cit., p. 163 ff.

§ *A general theory of linear sets*, loc. cit., p. 166 ff.

‖ Ibid., p. 167 ff.

¶ Ibid., p. 182 ff.

III. Linear systems and types of functions

In light of the results of the last section we will confine our attention to a system

$$(R, Q, V, F)$$

in which the elements are defined as follows.

R is either the rational number system with modulus ω or a field of integers modulo a prime p.

Q is a range; when Q is the range $1, 2, 3, \cdots, m$ we say that Q^m.

$V \equiv$ [all v on Q to R finitely non-zero], i. e., all functions (vectors) on Q to R such that there are only a finite number of elements of Q for which the corresponding functional values are $\neq 0$. Obviously if Q is finite $V =$ [all v on Q to R].

$F \equiv$ [all f on V to V], i. e., all functions on V to V.

The remainder of this paper studies the solutions of (A) when the x's are replaced by the v's. By ${}_nC_i\, v$ we shall mean $v+v+v+ \cdots$ (${}_nC_i$ terms) and by $(-1)^n$ the element 1 of R if n is even and -1 of R if n is odd.

A function f is of type A^k if f satisfies (A) when $n=k$ but does not satisfy (A) if $n<k$.

We proceed to define a class of functions in F which are analogous to polynomials.

Consider $H \equiv$ [all functions h, finitely non-zero, on Q to the class of integers $1, 2, 3, \cdots$]. We define the weight of any such h,

$$w(h) \equiv \sum_q h(q) \ .$$

We define, letting $0^0=1$,

$$v^h = \prod_q (v(q))^{h(q)} \quad (v,h) \ . \text{*}$$

This notation is of decided convenience due to the many theorems which it suggests, for instance

(i) $$v.).v^{0_H} = 1 \ .$$

Here by 0_H we mean the function in H which is identically zero.

(ii) $$h_1\, h_2\, v.).\ v^{h_1}v^{h_2} = v^{h_1+h_2} \ .$$

(iii) $$r\, v\, h\, .).(rv)^h = r^{w(h)}v^h \ .$$

Define

$${}_{h_1}C_h \equiv \prod_q {}_{h_1(q)}C_{h(q)}(h_1,h) \ ;$$

* This notation is essentially the same as that used by E. H. Moore in his paper entitled *On power series in general analysis*, Mathematische Annalen, vol. 86, p. 32 ff.

then

$$h_1\, v_1\, v_2 .).(v_1+v_2)^{h_1} = \sum_h {}_{h_1}C_h v^h v^{h_1-h} .$$

The symbol ${}_{h_1}C_h$ may be given a combinatory interpretation. Consider a class of $h_1(q)$ objects for each element of Q; then the number of ways $h(q)$ of these elements can be selected from each of the corresponding classes is the number ${}_{h_1}C_h$, of course, using the ordinary conventions that

$$n .). {}_nC_0 = 1 \text{ and } n_1 > n_2 .). {}_{n_2}C_{n_1} = 0 .$$

Clearly

$$w(h) > w(h_1) .). {}_{h_1}C_h = 0$$

and

$$h .). {}_hC_h = 1 .$$

We say that a function is of type P^n if it is of the form

$$\sum_h v_h v^h \quad (v)$$

and such that

(i) $$\exists h : \ni : w(h) = n \cdot v_h \neq 0_V ,$$

(ii) $$w(h) > n .). v_h = 0_V .$$

Definition of $m(h)$:

$$m(h) = r \mathrel{\vdots\equiv\vdots} r = \text{maximum value of } (h(q) \ (q)).$$

We say that f is of type P_n if it is a function of the form

$$\sum_h v_h v^h \quad (v)$$

and such that

(i) $$\exists h : \ni : m(h) = n \cdot v_h \neq 0_V ,$$

(ii) $$m(h) > n .). v_h = 0_V .$$

Definition of a linear m-fold:

V_0 is a linear m-fold $\vdots\equiv\vdots\ \exists v_{00}, (v_1, \cdots, v_m)$ linearly independent

$$.:\ni:. V_0 \supset v :\sim: \exists r_1 \cdot \cdot \cdot r_m$$

$$.\ni. v = v_{00} + \sum_i^{1.m} v_i r_i .$$

In such a case we say that $V_0 = \{v_{00}; v_1, \cdots, v_m\}$.

THEOREM 1. *If r belongs to R and v_1, v_2, and v_3 belong to a linear m-fold V_0, then $v_1 + r(v_2 - v_3)$ belongs to V_0.*

The proof is obvious.

THEOREM 2 (the converse of Theorem 1). *If for every r of R and v_1, v_2 of a set V_0, $v_1+r(v_1-v_2)$ belongs to V_0 and V_0 is contained in a linear m-fold, then V_0 is a linear m_1-fold where $m_1 \leqq m$.*

Proof. Any element of V_0 is effective as v_{00} and the $v_1, \cdots, v_m$ of the definition may be readily obtained by the use of the hypothesis. Call the greatest common subset of two sets of vectors their intersection.

THEOREM 3. *The intersection, if it exists, of a linear m_1-fold and a linear m_2-fold is a linear m_3-fold where m_3 is equal to or less than the lesser of m_1 and m_2; or it is a vector v.*

In light of Theorems 1 and 2 the proof is obvious.

Definition: A linear m_1-fold, V_1, and a linear m_2-fold, V_2, $m_1 \geqq m_2$, are said to be parallel provided the difference of every two vectors of V_2 is the difference of two vectors of V_1.

In this case either $V_1 \supset V_2$ or V_1 and V_2 have no intersection.

THEOREM 4.* *If Q is the finite range $(1, 2, 3, \cdots, m+1)$, if V_0 is a linear m-fold contained in the linear $(m+1)$-fold $\{0_V; \delta_1, \cdots, \delta_{m+1}\}$ (where δ_i is the function f for which $f(i)=1$ and $f(j)=0(i\neq j)$) then there exists uniquely a k and $r_0, \cdots, r_{k-1}$ such that $V_0=\{v_0; v_1, \cdots, v_m\}$ where*

$$v_0 = r_0\delta_k \ ,$$

$$v_j=\delta_j+r_j\delta_k \qquad (j<k),$$

$$v_j=\delta_{j+1} \qquad (j\geqq k).$$

THEOREM 5. *Relative to R with a modulus p, in every linear m-fold there exist $(p^{m+1}-p)/(p-1)$ distinct linear $(m-1)$-folds.*

Proof. For each determination of k in Theorem 4 there are p^k possible sets of values for the numbers r_0 to r_{k-1}. Hence the number of linear $(m-1)$-folds contained in the linear m-fold $\{0_V; \delta_1, \cdots, \delta_m\}$ is $\sum_{k=1}^{m} p^k = (p^{m+1}-p)/(p-1)$. Clearly in light of the definition of a linear m-fold and Theorems 1 and 2, this carries with it the theorem for the general case.

One should note here that relative to the matrix (r_{ij}) where i and j run from 1 to n, there exists for every set of vectors $v_j (j=1, \cdots, n)$ of V a set of vectors $v_i' (i=1, \cdots, n)$ such that

$$\sum_{i}^{1,n} v_i' r_{ij} = v_j \quad (j) \tag{1}$$

---

* *A general theory of linear sets*, loc. cit., see p. 185 ff.

if and only if the determinant $||r_{ij}||$ does not vanish. This follows at once from the analogous classical theorem, since condition (1) may be written

$$\sum_{i}^{1,n} v_i'(q) r_{ij} = v_j(q) \qquad (j,q) ,$$

which are purely numerical conditions and for every q the existence of the set $[v_i'(q)]$ $(i=1, \cdots, n)$ for every set $[v_j(q)]$ $(j=1, \cdots, n)$ is dependent on the non-vanishing of $||r_{ij}||$.

IV. The solution of (A). Part A

THEOREM 1. *If f is of type P^n then there exists a $k \leqq n$ such that f is of type A^k.*

Proof. Clearly the theorem follows from the proof of it for the special case in which f is of the form $(v_2 \, v^h(v))$, where $w(h) = n$. Clearly for any v_1 the first difference

$$v_2(v+v_1)^h - v_2 v^h$$
$$= v_2 \sum_{h_1}^{h_1 \pm h} {}_hC_{h_1} v^{h_1} v^{h-h_1} \quad (v)$$

is of the form $\sum_h v_h \, v^h \, (v)$, where $v_h \neq 0_V.).w(h) < n$. Hence the theorem is true for $n = n_1$ if it is true for $n = n_1 - 1$. However we know that it holds for the case $n = 0$, since in that case f is a constant and the first difference vanishes. Hence by induction the proof is complete.

THEOREM 2. *If f_1 and f_2 are each of the type A^k, $k \leqq n$, and for $n+1$ vectors of a linear 1-fold $V_0 = \{v_{00}; v_1\}$ the functional values of f_1 are equal to the functional values of f_2, then, as on V_0, f_1 is equal to f_2.*

Proof. Clearly since the $(n+1)$st difference vanishes Newton's interpolation formula $f(v_0 + lrv_1) = v_0 + l\Delta_1 + \frac{1}{2}l(l-1)\Delta_2 + \cdots$ holds for any set of values of the form $v_{00} + lrv_1$ for every integer l and any fixed value of r in R. Hence as on such a set f_1 is of the type P^{k_1} and f_2 is of the type P^{k_2}, where both k_1 and k_2 are equal to or less than n. Hence if $p < n+1$ the theorem is vacuously fulfilled. If $p = n+1$ the theorem is obvious. If $p > n+1$ or R has a modulus ω the well known theorem that a polynomial of degree equal to or less than n is determined by its values for $n+1$ values of the argument carries with it our theorem since both are equivalent to the non-vanishing of the same determinant.

Part B. Q finite and R with a modulus p

In this section we will consider Q the finite range $1, 2, \cdots, m$ and R with a modulus p.

In this section we order the elements of R linearly rather than cyclically. Thus $0<1<2 \cdots <p-1$.

THEOREM 1. *If two functions f_1 and f_2 are each of type P_{k_1} and P_{k_2} respectively $(k_1 \leqq k_2 \leqq n < p)$, and $V_1 =$ [all vectors of the form $\sum_{i=1}^{m} r_i \delta_i$ $(r_i \leqq n(i))$] and, as on V_1, $f_1 = f_2$, then $f_1 = f_2$ and corresponding coefficients are equal.*

Proof. This theorem is, of course, equivalent to the theorem that if f is of type P_k $(k \leqq n < p)$ and if f as on V_1 is identically 0 then f is identically 0_V and all the coefficients are 0_V.

Consider the coefficients of the polynominal in r_1 which are themselves polynomials in r_2 to r_m with coefficients in V which are of degree equal to or less than k in r_2 to r_m and vanish for all values of r_2 to r_m less than or equal to n, since the polynomial vanishes for $n+1$ values of r_1 whenever r_2 to r_m are all equal to or less than n; hence we have a reduction process and every coefficient must be 0_V.

THEOREM 2. *If $V_1 =$ [all vectors of the form $\sum_{i=1}^{m} r_i \delta_i (r_i \leqq n \ (i))$] where $n < p$, and g is a function on V_1 to V, then there exists one and only one function f of type $P_k (k \leqq n)$ such that f (as on V_1) $= g$.*

Proof. There are $p^{m(n+1)^m}$ distinct functions of type $P_k(k \leqq n)$ and since by Theorem 1 these functions as on V_1 are distinct they must as on V_1 be the $p^{m(n+1)^m}$ distinct functions on V_1 to V.

Hence every function f on V to V is of type P_k where $k<p$. Every function is, therefore, of type P^l where $l \leqq (p-1)m$.

THEOREM 3. *If f is of type $A^k (k < p-1)$ then f is of type P^k and conversely.*

Proof. f is of type P_l where $l<p$, i. e. f is of the form $\sum_h v_h v^h$ where $m(h) > p-1.).v_h = 0_V$. As on any linear 1-fold with typical element $v_0 + rv_1$, Newton's interpolation formula shows that f is a polynomial in r with coefficients in V of degree less than or equal to k. By substitution of $v = v_0 + rv_1$ it is seen that

$$\begin{aligned} f &= \sum_h v_h (v_0 + rv_1)^h \\ &= \sum_h v_h \sum_{h_1} {}_hC_{h_1} v_0^{h-h_1} r^{(h_1)} v_1^{h_1} \\ &= \sum_h v_h \sum_{k_1}^{w(h_1) \equiv k_1(p-1)} \sum_{h_1} {}_hC_{h_1} v_0^{h-h_1} r^{w(h_1)} v_1^{h_1} . \end{aligned}$$

Thus, since for any integer α, $r^\alpha = r^{\alpha+p-1}$,

$$p > k_1 > k .:):. v_0 \cdot v_1 :): \sum_h \sum_{h_1}^{w(h_1) \equiv (p-1)} v_{hh} C_{h_1} v_0^{h-h_1} v_1^{h1} = 0_V .$$

Hence, since this holds for all v_1 by Theorem 1,

$$p>k_1>k\ .:):.\ v_0\cdot w(h_1)\equiv k_1(p-1)\ :):\sum_h v_{hh}C_{h_1}v_0^{\,h-h_1}=0_V\ .$$

But this holds for every v_0, so

$$h\cdot p>k_1>k\ .:):.\ w(h_1)\equiv k_1(p-1)\ .).\ v_{hh}C_{h_1}=0_V\ .$$

Moreover,

$$p>k_1\cdot h:\ni:m(h)<p\cdot w(h)\geqq k_1\ .:):.\ \exists h_1:\ni:w(h_1)\equiv k_1(p-1)\cdot {}_hC_{h_1}\neq 0(p).$$

Hence,

$$p-1>k\cdot w(h)>k\ :):\ v_h=0_V\ .$$

Therefore, if f is of type A^k it is of type P^l where $l\leqq k$, but, if $l<k$, f would be of type A^{l_1} $(l_1\leqq l)$ (see Part A, Theorem 1) which is in contradiction to our hypothesis and, hence, f is of type P^k. The converse follows at once in light of Theorem 1 of Part A.

Part C. Q finite and R rational

In this part we will consider Q a finite range and R isomorphic with the field of all rational numbers.

In order to pass from the case where R has a modulus p to the case in which R is isomorphic with the field of rational numbers one should note the following theorem.

Theorem 1. *Relative to an m by m non-singular matrix $u=(r_{ij})$ whose elements are rational integers, and a set of integers c_i $(i=1,\cdots,m)$ there exists a proportionality factor λ and an integer n such that the solution of the equations*

$$\sum_j r_{ij}x_j=\lambda c_i\quad (i)$$

is the reduced solution of the congruences

$$\sum_j r_{ij}x_j\equiv\lambda c_i(p)\quad (i)$$

for every $p>n$.

By reduced solution is meant the solution for which each x is in absolute value equal to or less than $p/2$.

Proof. The system of linear equations

$$\sum_j r_{ij}x_j=c_i\|\mu\|\quad (i) \tag{1}$$

may be reduced by multiplication by integers (cofactors of the elements of $\|\mu\|$) and addition to the system

$$||\mu||x_j=\Delta_j||\mu|| \qquad (j=1,2,\cdots,m)$$

where Δ_j is the determinant of the matrix secured by replacing the jth column of μ by the c's. If n is greater than twice the absolute value of any of the multipliers or the coefficients of the unknowns in the equations of this reduction process, then for $p>n$ the reduction in the sense of a congruence modulo p would be formally identical. Since n is chosen greater in absolute value than $2||\mu||$, $2\Delta_j$ and $2\Delta_j||\mu||$, for every j, the reduced solution of the congruence

$$||\mu||x_j\equiv\Delta_j||\mu||(p) \text{ where } p>n$$

is $x_j=\Delta_j$.

Hence $||\mu||$ is effective as λ.

Theorem 2. *If on $V_1=$[all vectors of the form $\sum_{i=1}^m r_i\delta_i(0\leqq r_i\leqq n$ and integers) (i)] two functions f_1 and f_2 of type $A^k(k\leqq n)$ are equal, the two functions are equal.*

Proof. This can be readily arrived at by repeated application of Newton's interpolation formula. Obviously if f_1 is determined on V_1 it is determined for all vectors such that $v(i)$ is an integer or 0, less than or equal to n if $i\neq 1$ and unrestricted as to the value of $v(1)$. By another application the restriction on $v(2)$ may be removed, etc.

Moreover, from consideration of Newton's interpolation formula, if, as on V_1, f_1 has integral values for its coördinates, it has integral values for its coördinates for all P whose coördinates are integers.

Theorem 3. *If g is of type A^k $(k\leqq n)$ then there exists one and only one function f of type P_l $(l\leqq n)$ such that f (as on V_1) $=g$ (as on V_1).*

Proof. This theorem holds provided the $(n+1)^m$ by $(n+1)^m$ matrix of the system of equations for the coefficients of the terms of f

$$\sum_h v_h v^h=g(v) \qquad (v \text{ in } V_1)\ , \tag{1}$$

summation extending over all h such that $m(h)\leqq n$, is non-singular.

But this matrix reduced modulo $p\geqq n$ is non-singular since Theorem 2, Part B, holds.

Since g is of type A^k $(k\leqq n)$, there exists a λ such that $g_1\equiv\lambda g$ has integral coördinates for all vectors v whose coördinates are integers. If for all vectors v with integral coördinates we reduce the coördinates of $g_1(v)$ modulo p $(p>n)$ we have a new function g_1' (on all v with integral coördinates to all v with integral coördinates).

Moreover, g_1' is of type A^{k_1} $(k_1 \leqq k)$ and has the property that for every v_1 and v_2 with integral coördinates $g_1'(v_1+pv_2)=g_1'(v_1)$. Hence g_1' can be used to define a function of type A^{k_1} in the case where R has a modulus p. Moreover, if p is chosen sufficiently large, $k_1=k$, since there will exist some values of v_0 and v_1 relative to which equation (A) holds for $n=k$ but not for $n=k-1$ and if p is sufficiently large the same will be true for g_1'.

Hence by Theorem 1 and Theorems 2, 3, Part B, if p is sufficiently large the solution of equation (1) with g replaced by g_1' yields a function of type P^k and is the same as the solution of (1) with g replaced by g_1. But g_1 differs from g only by a constant factor. Hence f is of the type P^k.

But by Theorem 1, Part A, f is of type A^{k_2} $(k_2 \leqq k)$ and therefore by Theorem 2 of this part $f=g$. Hence we may conclude

THEOREM 4. *If V is the set of all vectors on a finite range Q to R then any function of type A^n is of type P^n if R is isomorphic with the field of all rational numbers or has a modulus $p>n+1$. If $p=n+1$ then f is of type P^k where $k \geqq n+1$ and P_l where $l<p$. There are no functions of type A^k where $k \geqq p$.*

Note: Another proof of the above theorem for the case in which R is rational exhibits a set of vectors such that for every corresponding set of values there exists one and only one function of type A^k $(k \leqq n)$ having for these vectors the assigned functional values. This proof is inserted here for the added insight it gives into the general problem of setting up solutions of equation (A) although as a proof of the above theorem it is much longer than the first one given.

Let ${}_nE_m$ be the number of ways that n or any integer less than n may be expressed as the sum of m integers or zeros, regarding each change of order as a new way. ${}_nE_m$ is also the number of distinct vectors h on the range $1, 2, 3, \cdots, m$ to the range $0, 1, 2, 3, \cdots$ such that $w(h) \leqq n$ and hence is the maximal number of non-zero terms in a polynomial of degree n in S_m, that is, the maximum number of non-zero terms in a function of type P^n when Q^m.

It is readily shown that

$$ {}_nE_m = \sum_i^{0\cdot n} {}_iE_{m-1} \, . $$

Consider a set E of ${}_nE_m$ vectors arranged in $n+1$ linear $(m-1)$-folds $S_0, S_1, \cdots, S_n$, such that every pair S_i, S_j $(i \neq j)$ intersect in a linear $(m-2)$-fold and these linear $(m-2)$-folds are pair by pair non-parallel, and such that no three intersect in the same $(m-2)$-fold. Let S_0 contain a set E_0 of ${}_nE_{m-1}$ vectors of E. Let S_1 contain a set E_1 of ${}_{n-1}E_{m-1}$ vectors of E not

in S_0. Let S_2 contain a set E_2 of ${}_{n-2}E_{m-1}$ vectors of E not in S_0 or S_1, etc. Moreover, let the ${}_{n-i}E_{m-1}$ vectors of E_i be arranged in $n-i+1$ linear $(m-2)$-folds $S_{i0}, \cdots, S_{i\,n-i}$ contained in S_i and not parallel to the intersection of any S_j $(j<i)$ with S_i and such that pair by pair they intersect in non-parallel linear $(m-3)$-folds and such that S_{i0} contains a set E_{i0} of ${}_{n-i}E_{m-2}$ vectors of E_i, S_{i1} contains a set E_{i1} of ${}_{n-i-1}E_{m-2}$ vectors of E_i not in S_{i0}, etc. Continue this mode of subdivision until all vectors of E are distributed in linear 1-folds.

The theorem that relative to any function g on E to V there exists one and only one function f of type P^k $(k\leqq n)$ such that, as on $E, f=g$, is equivalent to the non-vanishing of the determinant D of the matrix $\mu=(r_{vh})$ where $r_{vh}=v^h$ and v runs through the ${}_nE_m$ vectors of E while h runs through the ${}_nE_m$ vectors on the range 1, 2, 3, $\cdots$, m to the integers and 0 such that $w(h)\leqq n$.

Clearly by a reversible transformation S_0 may be taken as the linear $(m-1)$-fold $\{0_V;\delta_1,\cdots,\delta_{m-1}\}$. Hence, the minor corresponding to E_0 and the set H_0 of H such that, for any vector h_0 of H_0, $h_0(m)\neq 0$, is composed of zero elements only. Hence D is the product of the determinant D_1 of the minor μ_1 of μ which has elements r_{vh} such that $v\subset E_0$ and $h(m)=0$ and the determinant D_2 of the minor μ_2 which has elements r_{vh} such that v is not contained in E_0 and $h(m)\neq 0$.

The non-vanishing of D_1 is the condition that the theorem holds for $Q_1=1, 2, \cdots, m-1$.

$D_2=(\prod_v v(m))D_3$, where the product extends over all v in E not in E_0 and hence not in S_0 and hence is different from zero, whereas the non-vanishing of D_3 is equivalent to our theorem for the case of order $n-1$.

Hence by a finite number of steps we can reduce our theorem to the two special cases

(1) the case in which Q is singular;

(2) the case in which $n=0$.

In the first case the theorem is equivalent to the fact that a polynomial of degree less than n is determined by its value for $n+1$ values of the argument. In the second case f is a constant and the theorem is obvious.

If this work is to lead to an independent proof of Theorem 3 we should note that if two functions f_1 and f_2 are of type A^k $(k\leqq n)$, and, as on $S_0, \cdots, S_n, f_1=f_2$, then $f_1=f_2$. For consider any vector v not in any of the sets $S_0, \cdots, S_n$; then there exists a linear 1-fold V_0 containing v and $n+1$ distinct vectors $v_0, \cdots, v_n$ such that $S_i\supset v_i$. Hence, by Theorem 2, as on $V_0, f_1=f_2$ and hence $f_1(v)=f_2(v)$ for every v, i. e., $f_1=f_2$. However, the determination of a function of type A^k $(k\leqq n)$ on S_0 depends, in a similar manner, on the determination

of the function on $S_{00}, \cdots, S_{0n}$, the determination of the function on S_1 depends on its determination on $S_{10}, \cdots, S_{1\,n-1}$ and on the intersection of S_0 and S_1; its determination on S_2 depends on its determination on $S_{20}, \cdots, S_{2\,n-2}$ and the intersection of S_2 with S_1 and S_0; hence, we reduce the case for each linear m-fold considered to the case for $n+1$ linear $(m-1)$-folds and finally to the vectors of E. Hence any function of type A^k $(k \leqq n)$ must be identical with the function of type P^k having the same functional values on E.

This proof does not hold in general for the case where R has a modulus p a prime, as can be readily seen from the Theorem 5 of § III.

Part D. Q general

In this part we consider Q as a general range.

We shall show that any solution of (A) is of the form

$$\sum_h v_h v^h .$$

For any h_1 consider the set $Q_{h_1} = [\text{all } q \ni h_1(q) \neq 0]$; Q_{h_1} is finite. Consider Q_1 and Q_2, finite subsets of Q such that $Q_1 \supset Q_2 \supset Q_{h_1}$. By the theorem of the last section, f as on Q_1 is of the form

$$\sum_h v_{1h} v^h ,$$

and as on Q_2 is of the form

$$\sum_h v_{2h} v^h .$$

From the unique determination of the v's as shown in the above sections it is readily seen that v_{1h_1} (as on Q_2) $= v_{2h_1}$, since v_{1h_1} as on Q_2 must be completely determined by the value of f for v's for which $v(q) \neq 0$.). $Q_2 \supset q$. Let the set Q_2 swell.* It is readily seen that a v_{h_1} is determined uniquely as a limit. Consider, using the above determined v's, the function

$$\sum_h v_h v^h ;$$

we readily see that it is identical with f, since for any v all terms except those for which $Q_h \subset [\text{all } q \ni v(q) \neq 0]$ vanish and, for any such h, v_h as on Q_v is the same on any finite range containing Q_v as the coefficient we

* The consideration of this kind of limit has been given by E. H. Moore in his courses in General Analysis and a theory of limits including this as a special case has been discussed by E. H. Moore and H. L. Smith in their joint paper entitled *A general theory of limits*, American Journal of Mathematics, vol. 44 (1922).

would arrive at in the finite case. Moreover, if f is of the type A^n and R has a modulus ω or $p>n+1$ then f is of type P^n, but if $p=n+1$ we can only conclude that f is of type P_k $(k\leqq n)$ and that there exists an h such that $w(h)\geqq n$, $v_h\neq 0$.

V. Certain solutions directly in terms of a number system $\mathfrak{A}$ of type A

Returning to a number system $\mathfrak{A}$ of type A and considering functions on $\mathfrak{A}$ to $\mathfrak{A}$ the following theorem holds.

Theorem 1. *A function f of the form*

$$\sum_k^{0,n} a_{k0} \prod_i^{1,k} (aa_{ki}) \quad (a)\ , \tag{1}$$

where $\prod_{i=0}^n a_{ni}\neq 0$, is of the type A^l where $l\leqq n$.

Proof. Consider a function of the form

$$a_{k0} \prod_i^{1,k} (aa_{ki}) \quad (a), \text{ where } \prod_i^{0,k} a_{ki}\neq 0\ .$$

The first difference, for a difference a_1 in the argument, is

$$a_{k0}\prod_i^{1,k}(a+a_1)a_{ki} - a_{k0}\prod_i^{1,k}(aa_{ki}) \quad (a).$$

Clearly this function is of the form (1) where $n<k$, or is 0. Hence the theorem follows by induction.

VI. Continuous solutions

There is particular interest in the case where $\mathfrak{A}$ contains a field $\mathfrak{R}$ isomorphic with the real number system and commutative with $\mathfrak{A}$. It can be shown* that there exists a range Q such that $\mathfrak{A}$ is isomorphic with the set V of all finitely non-zero vectors on Q to $\mathfrak{R}$. If one defines the "écart" or distance of any two vectors v_1 and v_2, $\epsilon(v_1, v_2)$, as the maximum of the absolute value of the functional values of their difference, one can readily define (using the usual definition) a continuous function f on V to V thus:

$$f^C:\equiv: f:: \ni :: v\,, e>0\ .:):.\ \exists r_{ev}>0 :\ni: \epsilon(v, v_1)<r_{ev}$$
$$.).\epsilon\big(f(v), f(v_1)\big)<e\ .$$

Since by Newton's interpolation formula any solution of A is for any linear 1-fold $\{v_0\ ;\ v_1\}$ a polynomial in r as on all values of the form v_0+v_1r,

---

* *A general theory of linear sets*, loc cit., p. 182 ff.

where r is rational, then any continuous solution will be a polynomial as on $\{v_0 ; v_1\}$.

Moreover, in considering the case in which Q^m it is readily seen that any continuous solution is determined if the solution is determined for all v such that the functional values of v are rational. Hence, one may readily see from the results of the preceding sections

THEOREM 1. *Any continuous function of type A^n is a function of type P^n (where the terms are now understood to be relative to the Q, $\mathfrak{R}$, V system).*

To feel the force of the last theorem the following three corollaries are worth noting.

COROLLARY 1. *If $\mathfrak{A}$ is the real number system any continuous function of type A^n is a polynomial of degree n.*

This, of course, is obviously provable without reference to the theorem.

COROLLARY 2. *If $\mathfrak{A}$ is the complex number system any continuous function of type A^n is a polynomial of degree n in the coördinates x and y of $z=x+iy$, the coefficients of which are complex numbers.*

COROLLARY 3. *If $\mathfrak{A}$ is the system of real quaternions any continuous function of type A^n is a polynomial of degree n in the coördinates x, y, z, and w of $q=x+iy+jz+kw$, in which the coefficients are quaternions.*

UNIVERSITY OF WISCONSIN,
MADISON, WIS.

COMBINATORIAL ANALYSIS SITUS*

BY

J. W. ALEXANDER

1. **Introduction.** This is the first part of a paper on the topological theory of complexes. We begin by setting up a criterion for the homeomorphism of two complexes, expressed in terms of so-called *elementary transformations*, which are combinatorial in character. We then give a short account of the classical theory of connectivity and of a modified form of the theory where we operate with chains reduced modulo π. The theory of connectivity, modulo π, leads to connectivity numbers, modulo π, which are equivalent to the combined connectivity numbers and coefficients of torsion of the older theory. We close with strictly combinatorial proofs of the invariance of the connectivity numbers, modulo π.

The second part of the paper will be devoted to the theory of intersecting chains, out of which theory new topological invariants will be derived that cannot be expressed in terms of the classical ones.† Here, the advantage of operating with chains, modulo π, will become more apparent.

I. Combinatorial formulation

2. **Simplex and complex.** A *k-simplex* is the k-dimensional analogue of a tetrahedral region. Thus, for example, a *0-simplex* is a point, a *1-simplex* a line segment, a *2-simplex* a triangular plane region, and so on. The *boundary* of a k-simplex is made up of simplexes of dimensionalities $0, 1, \cdots, k-1$. It comprises $k+1$ *0-faces* (*vertices*), $\frac{1}{2}(k+1)k$ *1-faces* (*edges*), $\cdots$, $(k+1)!/(i+1)!(k-i)!$ *i-faces*, $\cdots$. For reasons of symmetry, we shall say that a k-simplex is its own k-face, though we shall not regard it as a member of its own boundary. A k-simplex is completely determined by its vertices. If we denote the latter by the marks $V_0, V_1, \cdots, V_k$ respectively, we may conveniently denote the simplex itself by the symbol

$$|V_0V_1 \cdots V_k|, \tag{2.1}$$

where the order of appearance of the marks V_i is immaterial.

For the purposes of this paper, a *complex* will be any finite set of simplexes such that (1) no two simplexes of the set have a point in common,

---

* Presented to the Society, February 28, 1925; received by the editors in March, 1925.

† Cf. abstracts in the Proceedings of the National Academy of Sciences, vol. 10 (1924), p. 99, p. 101, p. 493; vol. 11 (1925), p. 143.

(2) every face of a simplex of the set is itself a simplex of the set. This definition is less inclusive than the one ordinarily given.* However, there is no difficulty in showing that every complex of the more general sort is homeomorphic† with one of the restricted type defined above, so that, topologically speaking, no essential loss of generality results from thus narrowing down the definition.

The k-simplexes of a complex ($k=0, 1, 2, \cdots$) will be called the *cells* of the complex. 0-cells and 1-cells will also be called *vertices* and *edges* respectively. A complex is completely determined by those of its cells that are not on the boundaries of other cells; the presence of the remaining cells follows automatically from the definition of a complex. If the determining cells of a complex are all of the same dimensionality n, that is to say, if the complex consists entirely of n-cells and the cells on their boundaries, we shall call it an n-complex.

A complex may be represented schematically by the combined symbols of its determining cells and, therefore, ultimately, in terms of the symbols of its vertices. *We shall so arrange that all operations on complexes are ultimately expressible as combinatorial operations on the symbols associated with the vertices of the given complexes and of complexes derived from the latter by subdivision.*

A complex will be said to be *connected* provided that if V and V' are any two of its vertices there is always a sequence of vertices

$$V, V_1, V_2, \ldots V_i, V'$$

beginning with V and ending with V' such that every pair of consecutive vertices of the sequence is the boundary of an edge of the complex.

3. **Derived complexes.** Let

$$|V_0V_1 \cdots V_i| \qquad (i>0), \tag{3.1}$$

be an i-face of a k-simplex

$$|V_0V_1 \cdots V_iV_{i+1} \cdots V_k| \qquad (k\geqq i), \tag{3.2}$$

and let W be the center of mass of the face (3.1). We may then subdivide the simplex (3.2) into the $(i+1)$ k-simplexes

$$\begin{array}{c} | WV_1 \cdots V_iV_{i+1} \cdots V_k | \\ | V_0W \cdots V_iV_{i+1} \cdots V_k | \\ \cdots\cdots\cdots\cdots\cdots \\ | V_0V_1 \cdots WV_{i+1} \cdots V_k | \end{array} \tag{3.3}$$

* Cf., for example, Veblen, *Cambridge Colloquium Lectures*, Part I, *Analysis Situs*, p. 72. The symbol *VC* will hereafter be used to denote a reference to these Lectures.

† §4. *VC*, pp. 85, et seq.

together with certain separating faces of lower dimensionalities which need not be indicated explicitly. Furthermore, if the simplex (3 . 1) is a cell of a complex Φ, we may perform a subdivision of the type just described upon each cell (3 . 2) of the complex having the cell (3 . 1) as a face, whereby we shall transform the complex Φ into a new complex Ψ, made up of cells and sub-cells of Φ. The operation transforming the complex Φ into the complex Ψ will be called an *elementary subdivision*. It will be noticed that an elementary subdivision may be represented schematically in terms of vertex symbols, in conformity with the general plan of procedure laid out in § 2. For reasons of symmetry, we shall do away with the restrictions that the simplex (3 . 1) be of greater dimensionality than zero. When (3 . 1) is a 0-simplex, the operation of elementary subdivision becomes trivial and amounts merely to changing the name of the vertex V_0 from V_0 to W.

Now, suppose we assign an order to the cells of a complex Φ such that no cell precedes another of higher dimensionality than its own. Suppose, moreover, that we perform a series of elementary subdivisions, one corresponding to each cell of Φ, in such a manner that the new vertex W_i introduced in making the ith subdivision lies, in each case, at the center of mass of the ith cell of Φ (with respect to the ordering assigned to the cells of Φ). The effect of these combined operations is to transform the complex Φ into a new complex Φ_1, made up of proper sub-cells of Φ and having one vertex at the center of mass of each cell of Φ. We shall call the complex Φ_1 the *first derived complex* of Φ. The complex Φ_1 is independent of the exact ordering assigned to the cells of Φ so long as the condition is fulfilled that no cell shall be preceded by another of lower dimensionality than its own. The *nth derived complex* Φ_n of a complex Φ will be defined, by induction, as the first derived complex of the $(n-1)$st derived complex Φ_{n-1} of Φ. The derived complexes of Φ form an infinite sequence

$$\Phi_1 , \Phi_2 , \Phi_3 , \cdots$$

which will be called the *derived sequence* of Φ. We notice that the maximum diameter of the cells of Φ_i approaches zero as i increases without bound.

The vertices of the derived complexes of Φ will be called the *derived vertices* of Φ. They and their limit points make up the points of Φ.

4. **Homeomorphism in terms of derived vertices.** Two complexes are said to be *homeomorphic* provided there exists a one-one continuous correspondence between the points of one and the points of the other.

A necessary and sufficient condition that two compexes Φ and Ψ be homeomorphic is that there exist a one-one continuous correspondence between the derived vertices of Φ and of Ψ.

The sufficiency of the condition is obvious, because a one-one continuous correspondence between the derived vertices of Φ and Ψ induces a similar correspondence, by continuity, between all the points of Φ and of Ψ. To prove that the condition is necessary, we shall assume that there exists a one-one continuous correspondence K between the points of Φ and of Ψ and prove, in consequence, that there exists a similar correspondence between the derived vertices.

Let Φ_i be any derived complex of Φ. We shall first prove that there exists a correspondence K_1 differing from the given correspondence K by as little as we please and such that all the vertices of the complex Φ_i are mapped by K_1 upon vertices of some derived complex Ψ_j of Ψ. Consider the set of derived complexes Ψ_j of Ψ. Since the maximum diameter of the cells of a derived complex Ψ_j approaches zero as j increases without bound, it is possible to choose the index j in such a manner that the images with respect to K of no two vertices of the complex Φ_i lie on the same cell of the derived complex Ψ_{j-1} immediately preceding Ψ_j. Now, if we wish to form the complex Ψ_j from the complex Ψ_{j-1}, we must introduce a new vertex W_s at the center of each cell of Ψ_{j-1}. Instead of doing this, however, let us agree, whenever a cell of Ψ_{j-1} contains the image P_t of a vertex of Φ_i, to introduce the new vertex W_s at the point P_t rather than at the center of the cell in question. Then if we proceed, in other respects, just as we would if we were forming the complex Ψ_j, we are led to a complex Ψ'_j, similar to Ψ_j in cellular structure and deformable into Ψ_j by merely shifting the vertices W_s back to the centers of the respective cells of Ψ_{j-1} on which they lie. The deformation D undergone by the complex Ψ when Ψ'_{j-1} is transformed into Ψ_j clearly approaches zero as j increases without bound, since no point of Ψ ever leaves the cell of Ψ_{j-1} on which it lies initially. Since the correspondence K maps the vertices of the complex Φ_i upon vertices of the complex Ψ'_{j-1}, the correspondence $K_1 = KD$ maps the vertices of the complex Φ_i upon vertices of the complex Ψ_j and is, therefore, the correspondence we set out to find.

By an immediate induction, we now see that there exist an infinite sequence of complexes

$$(4.1) \qquad \Phi_{i_1}, \Psi_{i_2}, \Phi_{i_3}, \Psi_{i_4}, \cdots$$

and an infinite sequence of one-one continuous correspondences

$$(4.2) \qquad K_1, K_2, K_3, K_4, \cdots$$

such that

(i) the odd and even terms of (4.1) are sub-sequences of the derived sequences Φ and Ψ respectively;

(ii) the ith correspondence K_i maps the vertices of the $(i+1)$st complex of (4.1) upon vertices of the ith;

(iii) if the correspondence K_i pairs a derived vertex U of Φ with a derived vertex V of Ψ, all subsequent correspondences K_{i+s} pair the vertices U and V;

(iv) the sequence (4.2) converges uniformly to a limiting correspondence K_∞.

The limiting correspondence K_∞ is necessarily one-one continuous, by a well known theorem on uniformly convergent sequences. Moreover, it pairs the derived vertices of Φ with the derived vertices of Ψ in the manner required by the theorem.

5. **Elementary transformations.** If V is a vertex of a complex Φ we define the *star* $S(V)$ of Φ with center at V as the set of all cells of Φ having V as a vertex. Evidently, a 0-cell belongs to exactly one star of Φ, a 1-cell to two, a 2-cell to three, and so on. The *boundary* of a star will be the complex composed of the cells of Φ that are on the boundaries of cells of the star without themselves belonging to the star. Each i-cell of a star $(i>0)$ has exactly one $(i-1)$-face on the boundary of the star, and each $(i-1)$-cell on the boundary of the star is the $(i-1)$-face of exactly one i-cell of the star.

A set of vertices will be said to be *mutually adjacent* if, and only if, they are the vertices of a cell of Φ. It should be observed that a set of vertices may fail, as a whole, to be mutually adjacent in spite of the mutual adjacency of every pair of vertices in the set. Thus, for example, the three vertices of a triangle are not mutually adjacent with respect to the complex made up of the vertices and sides of the triangle, though they become mutually adjacent if the plane triangular region bounded by the triangle is added to the complex.

A necessary and sufficient condition that a set of vertices be mutually adjacent is that they be the centers of stars having at least one cell C^i in common. For if all their stars contain the cell C^i, they must themselves be vertices of C^i and, therefore, the vertices of some face of C^i. Conversely, if the vertices of the set determine a face of C^i, the cell C^i obviously belongs to all of their stars, from the very definition of a star. This simple condition for mutual adjacency will be useful in the sequel.

Now, consider any single-valued transformation τ carrying the vertices of a complex Φ into vertices of a complex Ψ in such a manner that sets of mutually adjacent vertices of Φ are carried into sets of mutually adjacent vertices of Ψ. From the definition of mutual adjacency, it follows that the vertices $V_0, V_1, \cdots, V_i$ of a cell of Φ are carried by τ into the vertices

$W_0, W_1, \cdots, W_i$ of a cell of Ψ. We may, therefore, extend the domain of definition of the transformation τ to the entire complex Φ by prescribing that it shall carry the interior and boundary of each cell

$$|V_0V_1 \cdots V_i| \tag{5.1}$$

of Φ by an affine transformation into the interior and boundary of the corresponding cell

$$|W_0W_1 \cdots W_i| \tag{5.2}$$

of Ψ, where the transformation is determined by the condition that V_0 be carried into W_0, V_1 into W_1, and so on. The extended transformation τ is single-valued and continuous over Φ; we shall call it an *elementary transformation.*

In defining an elementary transformation, *we do not wish to imply that distinct vertices of Φ are transformed into distinct vertices of Ψ.* In other words, it may very well happen that some of the vertices W_i in (5.2) coincide, and that the i-cell (5.1) of Φ is carried by a *degenerate* affine transformation (with vanishing determinant) into what we may call a *degenerate* i-cell (5.2) of Ψ,— properly speaking, a cell of dimensionality less than i. Thus, the inverse of the transformation τ need neither be single-valued nor everywhere defined over the complex Ψ. An elementary transformation is completely determined by its action on the vertices of a complex; consequently it is representable by a combinatorial operation on vertex symbols, again in accordance with our general program.

6. **Approximation of a continuous transformation.** Now, consider an arbitrary single-valued continuous transformation T of a complex Φ into a complex Ψ. (The inverse of the transformation T is not assumed to be single-valued or everywhere defined on Ψ.) We shall prove that there exists an infinite sequence of elementary transformations

$$\tau_1, \tau_2, \tau_3, \cdots \tag{6.1}$$

converging uniformly to the transformation T.

Let

$$\Phi_1, \Phi_2, \Phi_3, \cdots \tag{6.2}$$

and

$$\Psi_1, \Psi_2, \Psi_3, \cdots \tag{6.3}$$

be the derived sequences of Φ and Ψ respectively (§3). Then, if Ψ_i is any complex of the second sequence, there exists a complex Φ_j of the first such that the image with respect to the transformation T of each star of Φ_j

is completely covered by a star of Ψ_i. For the maximum diameter of a cell, and therefore of a star, of Φ_j approaches zero as j increases without bound; hence, the same is true of the maximum diameter of the image of a star of Φ_j, by elementary principles of uniform continuity. Thus, we see that there exists an infinite sub-sequence

$$\Phi_1', \Phi_2', \Phi_3', \cdots \qquad (\Phi_i' = \Phi_j) \tag{6.4}$$

of the sequence (6.2) such that, for every value of i, the image with respect to T of a star of Φ_i' is covered by a star of Ψ_i.

The transformation τ_i in (6.1) may now be determined by the condition that it shall carry the vertices of Φ_i' into vertices of Ψ_i in such a manner that the center of each star S^ϕ of Φ_i' is transformed into the center of one of the stars S^ψ of Ψ_i covering the image of the star S^ϕ with respect to the transformation T. To prove that this condition actually determines an elementary transformation, it is sufficient to show that mutually adjacent vertices of Φ_i' are carried into mutually adjacent vertices of Ψ_i. This follows at once, however, from the test for mutual adjacency given in § 5. For if a set of stars S_t^ϕ of Φ_i' $(t=1, 2, \cdots)$ have a cell C in common, so also must the corresponding stars S_t^ψ of Ψ_i covering the images of the stars S_t^ϕ respectively, since the stars S_t^ϕ all cover the image of the cell C. Finally, the transformations τ_i converge uniformly to the transformation T, because the image of a point P of Φ with respect to τ_i differs from the image of P with respect to T by less than the maximum diameter of a star of Ψ_i. But this maximum diameter approaches zero as i increases without bound. This completes the argument.

If the complexes Φ and Ψ are identical, the transformation T is a transformation of the complex Φ into itself. Now, it may happen that the transformation T is the identity. In this case, the approximating transformation τ_i carries each vertex V of the complex Φ_i' into the center of a star of the complex $\Phi_i(\Phi_i = \Psi_i)$, on which the vertex V lies, that is to say, into a vertex of the cell of Φ_i on which the vertex V lies. Such a transformation will be said to be *pseudo-identical*. We notice that it leaves fixed all the vertices of the derived complex Φ_i of Φ. Clearly, an infinite sequence of pseudo-identical transformations

$$\tau_{i_1}, \tau_{i_2}, \tau_{i_3}, \cdots \qquad (i_n > i_{n-1}),$$

converges uniformly to the identity, since the transformation τ_{i_n} displaces no point of the complex Φ by more than the maximum diameter of a cell of Φ_{i_n}.

7. **A condition for homeomorphism.** We are now ready to state in terms of elementary transformations a necessary and sufficient condition that two complexes Φ and Ψ be homeomorphic. To derive a necessary condition, let us assume that there exists a one-one continuous correspondence K between the points of Φ and the points of Ψ. Let

$$\Phi_1, \Phi_2, \Phi_3, \cdots \tag{7.1}$$

and

$$\Psi_1, \Psi_2, \Psi_3, \cdots \tag{7.2}$$

be the derived sequences of Φ and Ψ respectively. Then, if $\Delta_1 = \Phi_i$ is any complex of the first sequence, there exists a complex $\Delta_2 = \Psi_j$ of the second such that the image with respect to K of each star of Δ_2 is covered by a star of Δ_1 (cf. § 6). Hence, by induction, there exists an infinite sequence of complexes

$$\Delta_1, \Delta_2, \Delta_3, \cdots \tag{7.3}$$

with terms chosen alternately from (7.1) and (7.2), such that the image with respect to K of each star of each complex $\Delta_{i+1} (i>0)$ is covered by a star of the immediately preceding complex Δ_i. By the process of § 6, we may, therefore, determine an infinite sequence of elementary transformations

$$\tau_1, \tau_2, \tau_3, \cdots, \tag{7.4}$$

where the transformation τ_i carries the vertices of Δ_{i+1} into vertices of Δ_i and has the property that if a vertex V_{i+1} of Δ_{i+1} is carried into a vertex V_i of Δ_i the image with respect to K of the star of Δ_{i+1} with center at V_{i+1} is covered by the star of Δ_i with center at V_i.

Now, the transformations τ_{2i+1} of odd orders map the points of the complex Ψ upon points of the complex Φ and the transformations τ_{2i} of even orders map the points of the complex Φ upon points of the complex Ψ. Therefore, each product transformation of the form $\tau_{i+1}\tau_i$ maps one or the other of the complexes Ψ and Φ upon itself according as i is even or odd. Moreover, the transformation $\tau_{i+1}\tau_i$ carries the center V_{i+2} of a star S_{i+2} of Δ_{i+2} into the center of a star S_i of Δ_i such that S_i covers S_{i+2}. Therefore, since the center of a star is a vertex of all the cells of the star, the transformation $\tau_{i+1}\tau_i$ carries each vertex V_{i+2} of Δ_{i+2} into a vertex of the cell of Δ_i on which the vertex V_{i+2} lies. In other words, the transformation $\tau_{i+1}\tau_i$ is pseudo-identical.

The existence of the two sequences (7.3) *and* (7.4) *such that all product transformations of the form* $\tau_{i+1}\tau_i$ *are pseudo-identical is both a necessary and a sufficient condition that the complexes* Φ *and* Ψ *be homeomorphic.*

We have just seen that the condition is necessary; let us next show that it is sufficient. We prove, first of all, that the sequence

$$\tau_1\,,\tau_3\,,\tau_5\,,\ \cdot\ \cdot\ \cdot \tag{7.5}$$

consisting of the odd terms of (7.4) converges uniformly to a definite limiting transformation T of the points of Ψ into points of Φ. This is done by comparing two transformations τ_i and τ_j of (7.5) (i, j odd, $j>i$) with the auxiliary transformation

$$\tau_{ji}=\tau_j\tau_{j-1}\cdot\ \cdot\ \cdot\tau_{i+1}\tau_i\,.$$

This last transformation may be written in the form

$$\tau_{ji}=\tau_j(\tau_{j-1}\cdot\ \cdot\ \cdot\tau_i)\,,$$

where the expression in the parentheses is a pseudo-identical transformation of Φ carrying a point of a cell of Δ_i into a point on the boundary of the same cell. Therefore, the image of a point Q of Ψ with respect to τ_j differs from the image of the same point with respect to τ_{ji} by no more than the maximum diameter of a cell of Δ_i. But the transformation τ_{ji} may also be written

$$\tau_{ji}=(\tau_j\cdot\ \cdot\ \cdot\tau_{i+1})\tau_i$$

where, this time, the expression in the parentheses is a pseudo-identical transformation of Ψ carrying a point Q of a cell C of Δ_{i+1} into a point, or boundary point, of the same cell C. Hence, since the transformation τ_i carries the points and boundary points of the cell C of Δ_{i+1} into points or boundary points of a single cell of Δ_i, the image of Q with respect to τ_i also differs from the image of Q with respect to τ_{ji} by no more than the maximum diameter of a cell of Δ_i. By combining these two results, we see that the image of Q with respect to τ_i differs from the image of Q with respect to τ_j by no more than twice the maximum diameter of a cell of Δ_i, regardless of how j may be chosen. But the maximum diameter of a cell of Δ_i approaches zero as i increases without bound. Therefore, the uniform convergence of the sequence (7.5) is established. By a standard theorem on uniformly convergent sequences of continuous transformations, the limiting transformation T determined by (7.5) is also a continuous transformation. In a similar manner, we show that the sequence

$$\tau_2\,,\tau_4\,,\tau_6\,,\ \cdot\ \cdot\ \cdot \tag{7.6}$$

composed of the even terms of (7.4) converges uniformly to a continuous transformation U of the points of Φ into points of Ψ.

Now, the transformations T and U are, of course, single-valued, as well as continuous, over the complexes Ψ and Φ respectively. We know, more-

over, that $\tau_{i+1}\tau_i$ (for odd values of i) converges uniformly as i increases indefinitely, both to the identity and to UT. Therefore, we have

$$T^{-1}=U\ ,$$

which proves that the inverse of T is single-valued, as well as T itself. In other words, the transformation T determines a one-one continuous correspondence between the points of Φ and the points of Ψ. Hence, finally, the complexes Φ and Ψ are homeomorphic.

The above test for homeomorphism involves a limiting process. In proving the invariance of the topological constants, all that we shall actually need will be the following simpler, and purely combinatorial, theorem.

Let Φ and Ψ be homeomorphic complexes. Then, if Φ_i is any derived complex of Φ there exist derived complexes Ψ_j and Φ_k of Ψ and Φ_i, respectively, and elementary transformations τ and τ' such that

(i) *the transformation τ maps mutually adjacent vertices of Ψ_j upon mutually adjacent vertices of Φ_i;*

(ii) *the transformation τ' maps mutually adjacent vertices of Φ_k upon mutually adjacent vertices of Ψ_j;*

(iii) *the transformation $\tau'\tau$ is pseudo-identical.*

This theorem is contained in the previous one.

II. Connectivity

8. **Chains.** An *elementary i-chain* of a complex Φ will be defined as any symbolical expression of the form

$$\pm V_0V_1\cdots V_i\ , \tag{8.1}$$

where the marks $V_0, V_1, \cdots, V_i$ denote the vertices of an i-cell of Φ. In the expressions (8.1), we take account of the order in which the marks V_s are written, but make the convention that wherever two marks are permuted, the expression as a whole merely changes in sign. Thus, for example,

$$V_0V_1V_2=-V_1V_0V_2=V_1V_2V_0=\cdots\ .$$

It follows from this convention that there are two, and only two, essentially distinct elementary i-chains (8.1) associated with each i-cell $|V_0V_1\cdots V_i|$ of Φ.

Now, there are two distinct *orientations* on an i-cell $|V_0V_1\cdots V_i|$, just as there are two distinct directions on a line segment $|V_0V_1|$. It is therefore feasible to associate the symbol $V_0V_1\cdots V_i$ with the cell $|V_0V_1\cdots V_i|$ taken with one orientation and the symbol $-V_0V_1\cdots V_i$

with the same cell taken with the other. We prefer, however, to treat the expressions $\pm V_0V_1 \cdots V_i$ as purely symbolical, so as not to go into the question of just what is meant by an oriented cell. Let us merely remark, in passing, that the notion of oriented cells plays an important rôle in the theory of integration. If an i-cell is the domain of integration of a multiple integral, the sign of the integral depends upon the orientation of the i-cell.

Now, let the elementary i-chains (8.1) of the complex Φ be denoted by the abridged symbols

$$\pm \overset{i}{E_s} \qquad (s=1,2,\cdots,\alpha^i),$$

respectively. Then, any linear combination

$$K^i = x^1 \overset{i}{E_1} + x^2 \overset{i}{E_2} + \cdots + x^{\alpha^i} \overset{i}{E_{\alpha^i}} \tag{8.2}$$

of elementary i-chains $\overset{i}{E_s}$ with arbitrary integer coefficients x^s will be called an *i-chain* of Φ. We shall ordinarily write an expression such as (8.2) in the condensed form

$$K^i = x^s \overset{i}{E_s}, \tag{8.3}$$

where, by a convention similar to the one now current in tensor analysis, the double appearance in the same term of any variable index s, once as a subscript and once as a superscript, will be taken to imply a summation with respect to the index in question.

The i-chain K^i may be pictured as the domain of integration of an i-tuple integral, where the coefficients x^s denote that integration is to be performed $x^1, x^2, \cdots$ times over the oriented i-cells $\overset{i}{E_1}, \overset{i}{E_2}, \cdots$ respectively. However, we are not insisting on any specific interpretation of the symbol K^i.

A *chain* K of a complex Φ will be any linear combination of elementary i-chains of Φ, whether or not the elementary i-chains are all of the same dimensionality. Let us assign a distinctive symbol E_s to each elementary i-chain of Φ, irrespective of its dimensionality. A chain K will then be expressible by a relation of the form

$$K = x^s E_s, \tag{8.4}$$

where, according to our convention, summation with respect to the index s is to be understood.

A chain (8.4) will be said to *vanish*,

$$K = 0,$$

if, and only if, its coefficients x^s are all zero. The *negative* of the chain (8.4) will be the chain $-x^sE_s$. The *sum* and *difference* of two chains x^sE_s and

$y^s E_s$ will be the chains $(x^s+y^s)E_s$, and $(x^s-y^s)E_s$ respectively. The chains

$$K_t = x_t^s E_s \qquad (t=1, 2, \cdots, r)$$

will be said to be *linearly independent* if no linear combination of them with integer coefficients vanishes

$$\lambda^t K_t = 0$$

unless all the coefficients λ^t are equal to zero.

9. **Boundaries of chains; homologies.** The *boundary* of an elementary i-chain $V_0V_1 \cdots V_i$ will be defined as the $(i-1)$-chain

$$\sum_{s=0}^{i} (-1)^s V_0 \cdots V_{s-1}V_{s+1} \cdots V_i \,. \tag{9.1}$$

Thus, if we use the notation

$$K \to K'^*$$

to indicate that K is bounded by K', we shall have

$$\begin{aligned} &V_i \to 0 \,, \\ &V_iV_j \to V_j - V_i \,, \\ &V_iV_jV_k \to V_jV_k - V_iV_k + V_iV_j \,, \\ &\cdots\cdots\cdots\cdots\cdots \end{aligned} \tag{9.2}$$

In the notation of § 8, the boundaries of the elementary i-chains E_s^i of a complex Φ are determined by relations of the form

$$E_s^i \to [i]_s^t E_t^{i-1} \tag{9.3}$$

where each of the coefficients $[i]_s^t$ is equal to 0, 1, or -1.

An i-chain K^i is, by definition, a linear combination of elementary i-chains E_s^i,

$$K^i = x^s E_s^i \,; \tag{9.4}$$

we define its *boundary* as the corresponding linear combination of the boundaries of the chains E_s^i:

$$K^i \to x^s [i]_s^t E_t^{i-1} \,. \tag{9.5}$$

---

* Poincaré used the congruence symbol $\equiv$ in place of the arrow $\to$. The notation here adopted is perhaps less liable to cause confusion, and has the advantage of emphasizing the unsymmetrical character of the relation of bounding.

This definition is of the sort required for the correct formulation of the generalized Green-Stokes theorem which expresses certain i-tuple integrals over i-chains as $(i-1)$-tuple integrals over the boundaries of the i-chains.

A chain is always expressible as a sum of elementary i-chains, not necessarily of the same dimensionality. We define its *boundary* as the sum of the boundaries of its component elementary chains. It is clear that if

$$K_s \to K_s' \qquad (s=1,2\cdots,\sigma) \tag{9.6}$$

are bounding relations among chains, so also is any linear combination

$$\lambda^s K_s \to \lambda^s K_s'$$

of the relation (9.6). Moreover, if the coefficients μ^s and ν^s of a bounding relation

$$\mu^s K_s \to \nu^s L_s \tag{9.7}$$

have a common factor λ,

$$\mu^s = \lambda \bar{\mu}^s, \qquad \nu^s = \lambda \bar{\nu}^s,$$

it is legitimate to divide both members of (9.7) by the factor λ and write

$$\bar{\mu}^s K_s \to \bar{\nu}^s L_s .$$

A chain will be said to be *closed* (as distinguished from *open*, or *bounded*) if its boundary vanishes. The boundary (9.1) of an elementary i-chain $V_0V_1\cdots V_i$ is always closed. For the boundary (9.1) of the i-chain is itself bounded by

$$\begin{aligned}&\sum_{t<s}\sum(-1)^{s+t}V_0\cdots V_{t-1}V_{t+1}\cdots V_{s-1}V_{s+1}\cdots V_i\\ &+\sum_{t>s}\sum(-1)^{s+t-1}V_0\cdots V_{s-1}V_{s+1}\cdots V_{t-1}V_{t+1}\cdots V_i .\end{aligned} \tag{9.8}$$

But the two sums in (9.8) differ in sign, and in sign only, as we see by interchanging the names of the variables s and t in the second summation. More generally, the boundary of every chain is closed, since the boundary of a general chain K is the sum of the boundaries of the component elementary chains of K. According to our definition (cf. the first of relations (9.2)), every 0-chain is closed.* The boundary of an open i-chain is a closed $(i-1)$-chain.

* A different convention about the closure of 0-chains is also feasible, cf., for example, *VC*, p. 110. The convention adopted by Veblen amounts to replacing the relations $V_i \to 0$ in (9.2) by the symbolic relations $V \to 1$. The boundary of a general 0-chain $\mu^i V_i$ is then the sum of the coefficients μ^i of the chain, and the chain is closed if, and only if, this sum vanishes. Either convention has points in its favor, but the one adopted in this paper seems to be the more convenient one in the theory of intersecting chains.

A closed chain K of a complex Φ will be said to be ***bounding***, or ***homologous to zero***,

$$K \sim 0 \ ,^{*}$$

if it is the boundary of some open chain L of Φ. Two chains K and K', whether open or closed, will be said to be *homologous*,

$$K \sim K' \ ,$$

if their difference $K-K'$ is homologous to zero. Clearly, two chains with different boundaries can never be homologous.

Every linear combination of homologies

$$K_s \sim 0 \qquad (s=1, 2, \cdots, \sigma)$$

is an homology:

$$\lambda^s K_s \sim 0 \ . \tag{9.9}$$

It is important to notice, however, that the homology $\lambda K \sim 0$ does not necessarily imply the homology $K \sim 0$, for a multiple λK of a chain K may bound even in cases where the chain K itself does not. In other words, *it is not generally permissible to divide through the coefficients of an homology by a common factor.*

A set of closed chains K_s will be said to be *linearly independent with respect to bounding* if no linear combination $\lambda^s K_s$ of them bounds, (9.9), unless all the coefficients λ^s of the combination vanish. Since a multiple λK of a chain K may bound although the chain K itself does not, it may happen that the set composed of a single non-bounding chain K is not linearly independent with respect to bounding.

10. **Connectivity numbers.** The maximum number of closed i-chains of a complex Φ that are linearly independent with respect to bounding will be called the *ith connectivity number* of the complex Φ and will be denoted by the symbol P^i.† We shall prove, presently, that the number P^i is a topological invariant of the complex Φ.

It is easy to see that the 0th connectivity number P^0 of a complex Φ is essentially positive and equal to the number of non-overlapping connected complexes out of which the complex Φ is formed. For, by the second relation in (9.2), no 0-chain $x^s E_s$ can bound unless the sum of all its coefficients is zero; consequently no elementary 0-chain can bound. Furthermore, two elementary 0-chains E_1^0 and E_2^0 belonging to the same connected portion

* We are here using the terminology and notation of Poincaré.

† Except for an additive constant equal to unity, this is simply the ith Betti number of Poincaré, *VC*, p. 110.

of Φ are always linearly dependent on one another, since the points to which they correspond may be connected by a broken line made up of vertices and edges, and since this broken line determines a 1-chain bounded by a 0-chain of the form $\pm E_1^0 \pm E_2^0$. Thus, a complete set of linearly independent 0-chains consists of one elementary 0-chain from each connected portion of Φ.

The *connectivity* P of a complex Φ will be defined as the sum with respect to i of all the connectivity numbers P^i of Φ. The number P is equal to the maximum number of linearly independent chains of the complex Φ.

It is a simple matter to calculate the connectivity P of a complex Φ from a knowledge of the fundamental relations

$$E_s \to \mu_s^t E_t \qquad \left(\begin{matrix} s,t=1,2,\cdots,\alpha\,; \\ \mu_s^t=0,1,-1 \end{matrix}\right) \tag{10.1}$$

determining the boundaries of the elementary chains of Φ. Let ρ be the rank of the matrix of coefficients μ_s^t and α the number of elementary chains E_s. Every closed chain is the left-hand member of a linear combination of relations (10.1) such that the right-hand member vanishes; hence the number of independent closed chains is $\alpha-\rho$. Moreover, every homology among the closed chains is a linear combination of the homologies

$$\mu_s^t E_t \sim 0 \tag{10.2}$$

determined by the right-hand members of (10.1), since the boundary of an arbitrary chain K is equal to the sum of the boundaries of the elementary i-chains composing K. But the number of independent homologies (10.2) is equal to ρ. Therefore, finally, the maximum number of closed chains independent with respect to bounding is

$$P=\alpha-2\rho\,. \tag{10.3}$$

The individual connectivity numbers P^i may also be calculated in a similar manner. Let

$$\begin{matrix} E_s^{i+1} \to [i+1]_s^t E_t^i \\ E_t^i \to [i]_t^u E_u^{i-1} \end{matrix} \qquad \left(\begin{matrix} s=1,2,\cdots,\alpha^{i+1} \\ t=1,2,\cdots,\alpha^i \\ u=1,2,\cdots,\alpha^{i-1} \end{matrix}\right) \tag{10.4}$$

be the relations determining the boundaries of the elementary $(i+1)$- and i-chains of Φ respectively, and let ρ^{i+1} and ρ^i be the ranks of the matrices of coefficients $[i+1]_s^t$ and $[i]_t^u$ respectively. Then, by a similar argument to the last, we have

$$P^i = (\alpha^i - \rho^i) - \rho^{i+1}\,, \tag{10.5}$$

where, for $i=n$, we must put $\rho^{n+1}=0$. By eliminating the numbers ρ^i in relations (10.5) we obtain the so-called Euler-Poincaré formula,

$$\sum_{i=0}^{n} (-1)^i (P^i - \alpha^i) = 0\,. \tag{10.6}$$

The above derivation of formulas (10.5) and (10.6) is due to Poincaré. The simplified appearance of (10.6) is due to the modified definition of the numbers P^i that we have adopted.

Since the boundary of each elementary $(i+1)$-chain is closed, we also have

$$E_s^{i+1} \to [i+1]_s^t E_t^i \to [i+1]_s^t [i]_t^u E_u^{i-1} = 0\,,$$

whence, the product of the matrices of coefficients $[i+1]_s^t$ and $[i]_t^u$ must vanish:

$$[i+1]_s^t [i]_t^u = 0\,.\,* \tag{10.7}$$

For a similar reason, the square of the matrix of coefficients μ_s^t in (10.1) must also vanish,

$$\mu_s^t \mu_t^u = 0\,. \tag{10.8}$$

11. **Coefficients of torsion.** The elementary i-chains E_s^i of a complex Φ are a *minimal base* of the set of all i-chains of Φ. If a set of i-chains F_s^i is also a minimal base, the chains F_s^i are expressible in terms of the chains E_s^i by relations of the form

$$F_s^i = \gamma_s^t E_t^i \qquad (\gamma_s^t \text{ integers})\,, \tag{11.1}$$

and the chains E_s^i in terms of the chains F_s^i by relations of the form

$$E_t^i = \bar{\gamma}_t^u F_u^i \qquad (\bar{\gamma}_t^u \text{ integers})\,. \tag{11.2}$$

Moreover, since the product of the transformations (11.1) and (11.2) is the identity, the matrix of coefficients γ_s^t is the inverse of the matrix of coefficients $\bar{\gamma}_t^u$; i.e.,

$$\gamma_s^t \bar{\gamma}_t^u = \delta_s^u \qquad \left(\delta_s^u = \begin{cases} 0, & u \neq s \\ 1, & u = s \end{cases}\right).$$

It also follows from this that the product of the determinants $|\gamma_s^t|$ and $|\bar{\gamma}_t^u|$

* *VC*, p. 107.

is equal to unity. Therefore, the two determinants themselves are either both equal to $+1$ or both equal to -1.

For future reference, we recall the well known theorem that any transformation of the type (11.1) with determinant numerically equal to unity may be expressed as a product of *generating transformations* each of which is of one or the other of the following sorts:

(i) A transformation which leaves all but one of the members of a base invariant,

$$\overline{K}^i_s = K^i_s \qquad (s \neq j) \,,$$

and transforms the remaining member K^i_j by the simple addition of another member of the base

$$\overline{K}^i_j = K^i_j + K^i_k \qquad (k \neq j) \,.$$

(ii) A transformation which leaves all but one of the members of a base invariant,

$$\overline{K}^i_s = K^i_s \qquad (s \neq j),$$

and merely changes the sign of the remaining member,

$$\overline{K}^i_j = -K^i_j \,.$$

Suppose we change the base of the i-chains of Φ according to (11.1) and the base of the $(i-1)$-chains according to

$$E^{i-1}_u = \bar{\epsilon}^v_u F^{i-1}_v \,, \qquad |\bar{\epsilon}^v_u| = \pm 1 \,. \tag{11.3}$$

Then we see from (11.1) and (11.3) that the fundamental relations of bounding,

$$E^i_t \rightarrow [i]^u_t E^{i-1}_u \,, \tag{11.4}$$

transform into

$$F^i_s \rightarrow (i)^v_s F^{i-1}_v \,, \tag{11.5}$$

where we have

$$(i)^v_s = \gamma^t_s [i]^u_t \bar{\epsilon}^v_u \,. \tag{11.6}$$

Now, the rank of the determinant of the coefficients $[i]^u_t$ in (11.4) is equal to ρ^i; therefore, by generating transformations on the i-chains E^i_s we may transform (11.4) into a set of new relations with the property that the right-hand members of all but the first ρ^i of these last relations are equal to zero. In other words, there exists a minimal base for the i-chains

consisting of ρ^i open chains $\overset{i}{L}_t$ (such that every non-vanishing linear combination of the chains $\overset{i}{L}_t$ is open) and $\alpha^i - \rho^i$ closed chains $\overset{i}{K}_s$. Let us choose a base of this sort for each value of i. We may then write the fundamental relations of bounding (11.4) in the form

$$\begin{matrix} \overset{i}{L}_s \to (i)_s^t \overset{i-1}{K}_t \\ \\ \overset{i}{K}_u \to 0 \end{matrix} \qquad \begin{pmatrix} s = 1, 2, \cdots, \rho^i \\ t = 1, 2, \cdots, \alpha^{i-1} - \rho^{i-1} \\ u = 1, 2, \cdots, \alpha^i - \rho^i \end{pmatrix}. \tag{11.7}$$

We here use the fact, already noticed, that the boundary of an open i-chain is closed.

A further simplification of the first group of relations in (11.7) is possible. By generating transformation on the open i-chains $\overset{i}{L}_s$ and on the closed $(i-1)$-chains $\overset{i-1}{K}_t$, we may reduce the matrix of coefficient $(i)_s^t$ to a normal form such that all elements are zero except the ones along the main diagonal, and such that these last elements are the elementary divisors of the matrix. The elements

$$(i)_1^1, \ (i)_2^2, \ \cdots$$

along the main diagonal of the normalized matrix have the property that each is divisible by its predecessors. The numerical values of the elementary divisors are called by Poincaré the *coefficients of torsion* of the complex Φ They are topological invariants of Φ. In our present treatment, we shall replace the connectivity numbers and coefficients of torsion by other invariants to be considered in the following section.

12. **Chains, modulo π.*** Now, let π be any integer greater than unity. By a *chain K, modulo π,* we shall mean any linear combination

$$K = x^s E_s \qquad (\text{mod } \pi) \ , \tag{12.1}$$

of elementary chains E_s with integer coefficients x^s reduced modulo π. The *boundary* of the chain K, modulo π, will be the boundary of the corresponding chain of the non-modular type, with coefficients reduced modulo π. A chain K, modulo π, will be *closed,* if its boundary vanishes,

$$K \to 0 \qquad (\text{mod } \pi) \ ; \tag{12.2}$$

it will be *bounding*, or *homologous to zero,*

$$K \sim 0 \qquad (\text{mod } \pi) \ , \tag{12.3}$$

if it is the boundary of an open chain K', modulo π; and so on.

* Chains, modulo 2, were considered by Veblen and Alexander, *n-dimensional manifolds,* Annals of Mathematics, ser. 2, vol. 14 (1912–13), pp. 163–178. These chains have a particularly simple geometrical interpretation.

If the modulus π is a prime number, the discussion of § 10 obviously continues to hold when the term "chain" is reinterpreted to mean "chain, modulo π," for there is an almost perfect analogy between the theory of non-modular linear equations and the theory of linear equations to a prime modulus. Hence, when π is prime, we are led to certain new invariants $P^i(\pi)$ $(i=0, 1, \cdots, n)$, analogous to the connectivity numbers P^i. We shall call the number $P^i(\pi)$ the *ith connectivity number* of Φ, modulo π. It is equal to the maximum number of closed i-chains, modulo π, linearly independent with respect to bounding. The numbers $P^i(\pi)$ are given by formulas

$$(12.4) \qquad P^i(\pi)=\alpha^i-\rho^i(\pi)-\rho^{i+1}(\pi) \ ,$$

analogous to (10.5), where $\rho^i(\pi)$ and $\rho^{i+1}(\pi)$ are the ranks, modulo π, of the matrices of coefficients $[i]^t_s$ and $[i+1]^u_t$ respectively. Like the invariants P^i, the modular numbers $P^i(\pi)$ satisfy an Euler-Poincaré relation

$$(12.5) \qquad \sum_{1=0}^{n} (-1)^i[P^i(\pi)-\alpha^i]=0$$

analogous to (10.6). Their sum with respect to i gives the connectivity $P(\pi)$, modulo π, of the complex Φ.

As a matter of fact, it is fairly obvious that relations (12.4) and (12.5) may be derived even when the modulus π is not a prime, provided we put the proper interpretation on the rank $\rho(\pi)$, modulo π, of a matrix. It may be well, however, to say a few words about the theory of linear dependence to a general modulus π, since most treatments presuppose that π is a prime number.

A set of chains, modulo π,

$$(12.6) \qquad K_1 \ , K_2 \ , \ \cdot \ \cdot \ \cdot \ , K_\sigma \qquad \pmod{\pi} \ ,$$

will be *linearly independent* if no linear combination of them vanishes,

$$\lambda^s K_s=0 \qquad \pmod{\pi} \ ,$$

unless all the coefficients λ^s vanish, modulo π. It should be noticed that, when π is not a prime, i. e.,

$$(12.7) \qquad \pi=\lambda\tau \ ,$$

a set consisting of a single non-vanishing chain K need not be linearly independent. For the chain K may be the τth multiple of a chain K', in which case we have by (12.7)

$$\lambda K=\lambda\tau K'=\pi K'=0 \qquad \pmod{\pi}.$$

By a *generating transformation* of a set of chains (12.6), modulo π, we shall mean a transformation of either of the following two sorts:

(i) A transformation which leaves all but one of the chains K_s invariant and alters the remaining one by the simple addition of another chain in the the set

$$\begin{aligned} \overline{K}_s &= K_s \\ \overline{K}_j &= K_j + K_k \end{aligned} \qquad (\text{mod } \pi) \tag{12.8}$$

$$(s=1,2,\cdots,j-1,j+1,\cdots,\sigma\,;\, j\neq k)\,.$$

(ii) A transformation which leaves all but one of the chains K_s invariant and multiplies the remaining one by an integer κ, such that the integer κ and modulus π are mutually prime,

$$\begin{aligned} \overline{K}_s &= K_s \\ \overline{K}_j &= \kappa K_j \end{aligned} \qquad (\text{mod } \pi) \tag{12.9}$$

$$(s=1,2,\cdots,j-1,j+1,\cdots,\sigma\,;\, j\neq k)\,.$$

Every generating transformation has an inverse which is expressible as a product of generating transformations. The inverse of (12.8) is

$$\begin{aligned} K_s &= \overline{K}_s \\ K_j &= \overline{K}_j - \overline{K}_s \end{aligned} \qquad (\text{mod } \pi)$$

$$(s=1,2,\cdots,j-1,j+1,\cdots,\sigma\,;\, j\neq k)\,,$$

while the inverse of (12.9) is

$$\begin{aligned} K_s &= \overline{K}_s \\ K_j &= \lambda \overline{K}_j \end{aligned} \qquad (\text{mod } \pi)\,,$$

where the coefficient λ is chosen to satisfy the diophantine equation

$$\lambda\kappa + \mu\pi = 1\,.$$

Since the integers κ and π are mutually prime, this last equation determines the integer λ uniquely.

THEOREM. *Let*

$$K = \lambda^s K_s \qquad (\text{mod } \pi)$$

be a linear combination of chains K_s, modulo π, and

$$\overline{K} = \lambda^s \overline{K}_s = \mu^s K_s \qquad (\text{mod } \pi)$$

the transform of K under a generating transformation of the chains K_s. Then, if the coefficients λ^s do not all vanish, modulo π, neither do the coefficients μ^s.

The theorem follows at once from the definition of a generating transformation. It would cease to be true if the constant κ in a generating transformation of the second sort were allowed to have a factor in common with the modulus π.

COROLLARY. *A generating transformation carries a linearly independent set into a linearly independent set.*

For a relation $\lambda^s\overline{K}_s=0 \pmod{\pi}$ would imply a relation $\mu^s K_s=0 \pmod{\pi}$.

Two sets of chains K_s and L_s, modulo π, will be said to be *equivalent* if the set of all chains linearly dependent on the chains K_s is the same as the set of all chains linearly dependent on the chains L_s.

THEOREM. *A necessary and sufficient condition that two linearly independent sets of chains, modulo π, be equivalent is that either be transformable into the other by one or more generating transformations.*

The sufficiency of the condition is obvious. To prove its necessity, let K_s $(s=1, 2, \cdots, \sigma)$ and L_t $(t=1, 2, \cdots, \tau)$ be two linearly independent sets of chains, modulo π, and let the notation be so chosen that σ is at least as great as τ. Then, since the two sets are equivalent, each chain K_s is a linear combination of chains L_t,

$$K_s=\lambda_s^t L_t \pmod{\pi}. \tag{12.10}$$

Now, we know by the general theory of matrices that if the chains K_s and L_s were not reduced modulo π, it would be possible, by generating transformations on the K_s's and on the L_s's, to reduce relations (12.10) to the form

$$K_s'=\mu_s^t L_t', \tag{12.11}$$

where all of the coefficients μ_s^t would vanish except the ones along the main diagonal of the matrix $\|\mu_s^t\|$. In the modular case, the corresponding generating transformations will therefore also reduce (12.10) to the form (12.11). In the modular case, we notice, further, that each of the coefficients μ_i^i along the main diagonal must be a prime residue, modulo π, for if we had

$$\mu_i^i q+\pi r=0 \qquad \text{(not summed for } i\text{)},$$

we would have

$$qK_i'=\mu_i^i qL_i'=0 \qquad \text{(not summed for } i\text{)},$$

whereas the chains K_i' are linearly independent, by the corollary to the last theorem. In particular, none of the coefficients μ_i^i can vanish; therefore, there must be exactly as many chains K_i' as there are chains L_i'. Finally, if all the coefficients μ_i^i along the main diagonal are prime residues, modulo π, the chains L_i' may be transformed into the chains K_i' by generating transformations of the second sort. It is sufficient, in fact, to multiply each chain L_i' by the inverse, modulo π, of the corresponding coefficient μ_i^i. Thus, the K_s's are equivalent to the K_s''s which are equivalent to the L_s''s, which are equivalent to the L_s's.

On the basis of these theorems, it follows immediately that when we are reducing to any modulus π, the fundamental relations of boundary

$$E_t^i \to [i]_t^u E_u^{i-1} \qquad (\mathrm{mod}\ \pi)$$

may be reduced to a normal form

$$\begin{gathered} L_s^i \to (i)_s^t K_t^{i-1} \\ K^i \to 0 \\ \begin{pmatrix} s=1,2,\cdots,\rho^i(\pi) \\ t=1,2,\cdots,\alpha^{i-1}-\rho^{i-1}(\pi) \\ u=1,2,\cdots,\alpha^i-\rho^i(\pi) \end{pmatrix}, \end{gathered} \qquad (\mathrm{mod}\ \pi) \tag{12.12}$$

analogous to (11.7), where all the coefficients $(i)_s^t$ vanish except those along the main diagonal, which last are not congruent to zero. The number of non-vanishing coefficients $(i)_s^s$ will be, by definition, the rank $\rho^i(\pi)$, modulo π, of the matrix $[i]_t^u$. If π is a prime, this definition is equivalent to the ordinary one.

If π is a prime, relations (12.12) may be reduced to such a form that all the diagonal elements $(i)_s^s$ are equal to unity. If π is not a prime, they may be reduced to such a form that each is exactly divisible by its predecessors (in the strict, non-modular sense) and that each is a factor of π. Thus, when π is not a prime, we obtain numbers analogous to the coefficients of torsion (§ 11).

The modular connectivity numbers $P^i(\pi)$ give exactly the same information about the complex Φ as the combined connectivity numbers P^i and coefficients of torsion. For, to derive the normalized modular relations of bounding (12.12) from the normalized non-modular ones (11.7), we must replace by a zero every coefficient of torsion in (11.7) containing the factor π, that is to say, every coefficient of torsion greater than or equal to π. Therefore, if $\sigma^i(\pi)$ is the number of coefficients of torsion of the ith order greater

than or equal to π, the rank of the matrix of coefficients $[i]^u_t$ is related to the rank, modulo π, of the same matrix in the following manner:

$$\rho^i = \rho^i(\pi) + \sigma^i(\pi) \ . \tag{12.13}$$

Consequently, by (10.5), (12.4), and (12.13), the modular connectivity number $P^i(\pi)$ is expressible in terms of the non-modular connectivity number P^i and coefficients of torsion according to the formula

$$P^i(\pi) = P^i + \sigma^i(\pi) + \sigma^{i+1}(\pi) \qquad [\sigma^0(\pi) = 0] \ . \tag{12.14}$$

On the other hand, the relations (12.14) may be solved successively for the numbers $\sigma^i(\pi)$ in terms of the connectivity numbers P^i and modular connectivity numbers $P^i(\pi)$. Moreover, for any particular complex Φ, there exists a positive integer π_0 such that

$$\rho^i(\pi) = \rho^i \qquad (\pi > \pi_0) \ ,$$

and hence,

$$P^i(\pi) = P^i \qquad (\pi > \pi_0 \ .) \tag{12.15}$$

Therefore, the connectivity numbers P^i and coefficients of torsion are functions of the modular connectivity number $P^i(\pi)$. Of course, it is never necessary to evaluate the invariants $P^i(\pi)$ for more than a finite number of values of π, in view of (12.15). While the theory of modular chains leads us to invariants $P^i(\pi)$ which are new in form only, we shall find it very serviceable in the developments to be treated in the second part of this paper.

III. Invariance of the topological constants

13. **Transformations of chains.** A *degenerate elementary i-chain* of a complex Φ will be defined as any symbol of the form

$$\pm V_0 V_1 \cdot \cdot \cdot V_i \tag{13.1}$$

containing at least one repeated vertex, where the vertices $V_0, V_1, \cdots, V_i$ are mutually adjacent vertices of Φ (§ 5). The *boundary* of a degenerate elementary i-chain will be defined according to the same law, (9.1), as the boundary of an ordinary elementary i-chain. A *generalized chain* will be any linear combination of elementary chains, whether ordinary or degenerate. Its boundary will be the sum of the boundaries of its component elementary chains.

Every elementary transformation τ (§ 5) carrying the vertices V_k of the complex Φ into vertices W_k of a complex Ψ may be thought of as carrying each elementary i-chain (13.1), whether ordinary or degenerate, into the elementary i-chain

$$\pm W_0 W_1 \cdot \cdot \cdot W_i \ . \tag{13.2}$$

It, therefore, carries every generalized chain of Φ into a generalized chain of Ψ. Moreover, if two generalized chains K and L of Φ are in the relation

$$K \rightarrow L \, , \tag{13.3}$$

they are carried respectively into chains K' and L' of Ψ such that

$$K' \rightarrow L' \, . \tag{13.4}$$

Now, we notice that the boundary of a degenerate elementary chain is a linear combination of degenerate elementary chains, just as the boundary of a non-degenerate elementary chain is a linear combination of non-degenerate elementary chains. For if a degenerate i-chain E^i with the boundary (9.1) contains more than two coincident vertices, every term in (9.1) contains coincident vertices, while if the chain E^i contains exactly two coincident vertices all but two of the terms in (9.1) contain coincident vertices, and the two which do not are of opposite signs, but otherwise identical, so that they cancel one another. Let us agree that the marks representing degenerate elementary chains are to be set equal to zero whenever they appear in the symbol for a generalized chain. Then, with this convention, a generalized chain reduces to an ordinary chain. Moreover, a relation of bounding, such as (13.3), is preserved under this reduction, for, by the remarks just made, the degenerate and non-degenerate portions of the left-hand member of (13.3) are bounded by the degenerate and non-degenerate portions of the right-hand member respectively. Thus, we may say that (13.3), regarded as a relation between ordinary chains, is carried by the transformation τ into (13.4), likewise regarded as a relation between ordinary chains. We conclude, therefore, that *an elementary transformation τ carries a closed chain into a closed (possibly vanishing) chain and a bounding chain into a bounding (possibly vanishing) chain.* Needless to say, however, the transformation τ may carry an open chain into a closed chain or a non-bounding chain into a bounding one.

It may be remarked, in passing, that the convention of setting degenerate elementary i-chains equal to zero is consistent with the notation of § 8, according to which an elementary chain $V_0V_1 \cdots V_i$ changes sign when any two of its vertices V_s are permuted.

14. **Invariance under elementary subdivisions.** Let ϵ be an elementary subdivision (§ 5) transforming a complex Φ into a complex Φ'. Then, as we recall, the subdivision ϵ is effected by introducing a new vertex W at the center of some i-cell

$$|V_0V_1 \cdots V_i| \tag{14.1}$$

of the complex Φ and subdividing each k-cell

$$|V_0V_1 \cdots V_iV_{i+1} \cdots V_k| \qquad (k \geqq i) \tag{14.2}$$

with the face (14.1) into the $i+1$ k-cells

$$|V_0 \cdots V_{t-1}WV_{t+1} \cdots V_iV_{i+1} \cdots V_k| \qquad (0 \leqq t \leqq i)\ , \tag{14.3}$$

together with certain cells of lower dimensionalities.

Now, let us agree to express each of the elementary chains of Φ corresponding to the cells (14.2) in terms of elementary chains of Φ' corresponding to the cells (14.3), in accordance with the following formulas:

$$\begin{aligned} & V_0 \cdots V_iV_{i+1} \cdots V_k \\ &= \sum_{t=0}^{i} V_0 \cdots V_{t-1}WV_{t+1} \cdots V_iV_{i+1} \cdots V_k \\ &= \sum_{t=0}^{i} (-1)^t WV_0 \cdots V_{t-1}V_{t+1} \cdots V_iV_{i+1} \cdots V_k\ . \end{aligned} \tag{14.4}$$

Since the remaining elementary chains of Φ are also elementary chains of Φ', it follows that the relations (14.4) enable us to express every chain K of Φ as a chain K' of Φ'.

LEMMA 1. *If the relations* (14.4) *identify a chain K of Φ with a chain K' of Φ', they also identify the boundary of the chain K with the boundary of the chain K'.*

To prove the lemma, it will be sufficient to show that the boundary of the left-hand member of each relation (14.4) may be identified with the boundary of the corresponding right-hand member (with the help of the other relations of the set); for a general chain K of Φ is merely a sum of elementary chains. The identification of the boundaries of the two members of (14.4) is self-evident geometrically. The formal proof is as follows.

The boundary of the left-hand member of (14.4) is

$$\sum_{t=0}^{k} (-1)^t V_0 \cdots V_{t-1}V_{t+1} \cdots V_k$$

which may be broken up into the two partial sums

$$\begin{aligned} & \sum_{t=0}^{i} (-1)^t V_0 \cdots V_{t-1}V_{t+1} \cdots V_k \\ & \qquad + \sum_{s=i+1}^{k} (-1)^s V_0 \cdots V_{s-1}V_{s+1} \cdots V_k\ . \end{aligned} \tag{14.5}$$

The boundary of the right-hand member is

$$
\begin{aligned}
&\sum_{t=0}^{i}(-1)^t V_0 \cdots V_{t-1}V_{t+1}\cdots V_k \\
&+\sum_{s<t}\sum_{t=0}^{i}(-1)^{t+s+1} W V_0 \cdots V_{s-1}V_{s+1}\cdots V_{t-1}V_{t+1}\cdots V_k \\
&+\sum_{t<s}\sum_{s=0}^{i}(-1)^{t+s} W V_0 \cdots V_{t-1}V_{t+1}\cdots V_{s-1}V_{s+1}\cdots V_k \\
&+\sum_{t=0}^{i}\sum_{s=i+1}^{k}(-1)^{t+s} W V_0 \cdots V_{t-1}V_{t+1}\cdots V_{s-1}V_{s+1}\cdots V_k\,.
\end{aligned}
\tag{14.6}
$$

But the second and third sums in (14.6) cancel one another, as may be seen by interchanging the names of the variables s and t in the second. Moreover, the first sum in (14.6) is identical with the first in (14.5) and the fourth sum in (14.6) with the second in (14.5) when this last sum is expressed as a chain of Φ', by means of relations (14.4). Thus, we have identified the boundaries of the left and right-hand members of each relation (14.4), and thereby proved the lemma.

In the notation that we have just been using, let $|V_0\cdots V_k|$ be a k-cell of the complex Φ such that the new vertex W lies at the center of an i-face $|V_0\cdots V_i|$ of $|V_0\cdots V_k|$. Properly speaking, there exists no elementary $(k+1)$-chain

$$
W V_0 \cdots V_k\,,
\tag{14.7}
$$

since the points $W, V_0, \cdots, V_k$ cannot be the vertices of a $(k+1)$-cell. For symmetry of expression, however, we shall speak of (14.7) as a *symbolic* $(k+1)$-chain. The *boundary* of the symbolic chain (14.7), formed according to (9.1), is

$$
\begin{aligned}
V_0\cdots V_k - &\sum_{t=0}^{i}(-1)^t W V_0\cdots V_{t-1}V_{t+1}\cdots V_k \\
&-\sum_{t=i+1}^{k}(-1)^t W V_0\cdots V_{t-1}V_{t+1}\cdots V_k\,,
\end{aligned}
$$

where the first two terms are made up of ordinary k-chains, while the last is made up of symbolic ones. Moreover, by (14.4), the first two terms cancel one another, so that the boundary of the symbolic $(k+1)$-chain (14.7) reduces to a sum of symbolic k-chains,

$$
\sum_{t=i+1}^{k}(-1)^t W V_0\cdots V_{t-1}V_{t+1}\cdots V_k\,.
\tag{14.8}
$$

In view of this fact, we may operate with symbolic chains in just the same way as with degenerate chains, by regarding them merely as symbols for zero.

A chain K will be said to *meet* a vertex V if the symbol $\pm V_0V_1\cdots V_i$ of one of its component elementary chains contains the mark V.

LEMMA 2. *Every chain K' of Φ' such that its boundary does not meet the new vertex W is homologous with a chain K of Φ.*

This proposition is also self-evident geometrically. The detailed proof is as follows:

We break up the chain K' into a pair of chains J' and L,

$$K' = J' + L\,, \tag{14.9}$$

composed respectively of the elementary chains of K' that meet the new vertex and of those that do not. The chain J' is of the form

$$J' = \sum \pm WV_{s_1}\cdots V_{s_q}\,; \tag{14.10}$$

the boundary of J' is of the form

$$\begin{aligned} B = {} & \sum \pm V_{s_1}\cdots V_{s_q} \\ & - \sum_{t=1}^{q}\sum \pm(-1)^t WV_{s_1}\cdots V_{s_{t-1}}V_{s_{t+1}}\cdots V_{s_q}\,. \end{aligned} \tag{14.11}$$

Moreover, since the boundaries of K' and L do not meet the vertex W, neither does the boundary B of J'. Therefore, the double sum in (14.11) must vanish by mutual cancellation of its terms,

$$\sum_{t=1}^{q}\sum \pm(-1)^t WV_{s_1}\cdots V_{s_{t-1}}V_{s_{t+1}}\cdots V_{s_q} = 0 \tag{14.12}$$

leaving the relation

$$B = \sum \pm V_{s_1}\cdots V_{s_q}\,. \tag{14.13}$$

Now, let V_0 be any vertex of the cell of the complex Φ on which the new vertex W lies. Consider the sum of elementary chains

$$\sum \pm V_0WV_{s_1}\cdots V_{s_q} \tag{14.14}$$

(actual, degenerate, or symbolical) obtained by prefixing the mark V_0 to each elementary chain in J', (14.10). The boundary of (14.14) is

$$\begin{aligned} & \sum \pm WV_{s_1}\cdots V_{s_q} - \sum \pm V_0V_{s_1}\cdots V_{s_q} \\ & - \sum\sum \pm(-1)^t V_0WV_{s_1}\cdots V_{s_{t-1}}V_{s_{t+1}}V_{s_q} \sim 0\,. \end{aligned} \tag{14.15}$$

But the double sum in (14.15) must vanish by mutual cancellation of its terms, since it differs from the double sum in (14.12) merely by the presence of the mark V_0 at the head of each term. Therefore, (14.15) reduces to

$$(14.16) \qquad J' - M \sim 0 \qquad (M = \sum \pm V_0 V s_1 \cdot \cdot \cdot V s_q) .$$

From (14.9) and (14.16), we finally obtain

$$K' \sim M + L .$$

But the chains M and L do not meet the vertex W. Therefore, $M+L$ is a chain of the complex Φ, which proves the lemma.

15. **Invariance of the topological constant.** Let Φ and Ψ be two homeomorphic complexes. Then, by the corollary at the end of § 7, there exist derived complexes Ψ_j and Φ_k of Ψ and Φ respectively and elementary transformations τ and τ' such that (1) the transformation τ carries vertices of Ψ_j into vertices of Φ, (2) the transformation τ' carries vertices of Φ_k into vertices of Ψ_j, (3) the transformation $\tau'\tau$ is pseudo-identical.

Lemma. *Every i-chain K of the complex Φ, when expressed as a chain of Φ_k and operated on by the pseudo-identical transformation $\tau'\tau$, is left invariant by the transformation $\tau'\tau$.*

The proof is made by induction with respect to i. If i is zero, the lemma is immediate, since the pseudo-identical transformation $\tau'\tau$ leaves invariant the vertices of Φ. Let us, therefore, assume that the lemma has been proved for all k-chains of dimensionalities k less than i and show, first of all, that it is true for an elementary i-chain E^i corresponding to an i-cell C^i of Φ. We shall denote the boundary of E^i by K^{i-1},

$$E^i \rightarrow K^{i-1} .$$

Now, the chain E^i, when expressed as a chain of Φ_k, consists of a sum of elementary i-chains corresponding to sub-cells of C^i. Moreover, each vertex of each of these sub-cells is carried by the transformation $\tau'\tau$ into a vertex of C^i; therefore, E^i must be carried into some multiple λE^i of itself, such that

$$\lambda E^i \rightarrow \lambda K^{i-1} .$$

However, by the hypothesis of the induction, the boundary K^{i-1} of E^i is carried into itself. Therefore, $\lambda = 1$. Therefore, finally, E^i is carried into itself. It follows at once that every chain K of Φ is left invariant by the transformation $\tau'\tau$ since each of the component elementary chains of K is left invariant.

We may now complete the proof of the invariance of the connectivity numbers in a few lines. Let

$$K_s^i \qquad (s=1, 2, \cdots, P^i)$$

be a complete set of closed i-chains of the complex Φ, linearly independent with respect to bounding. These chains, expressed as chains of Φ_k, are carried by the transformation τ' into chains

$$L_s^i, \qquad (s=1, 2, \cdots, P^i)$$

of Ψ_j respectively. Moreover, the chains L_s^i are carried back again into the chains K_s^i respectively by the transformation τ, in consequence of the lemma. It follows that the chains L_s^i must be linearly independent with respect to bounding; for a relation of bounding

$$\lambda^s L_s^i \sim 0$$

between the chains L_s^i would go over into a relation of bounding

$$\lambda^s K_s^i \sim 0$$

between the chains of K_s^i, contrary to hypothesis. Thus the connectivity number P^i must be at least as great for the complex Ψ as for the complex Φ. But, by reversing the argument, the number P^i must be at least as great for the complex Φ as for the complex Ψ. Hence, finally, the number P^i must be the same for the two complexes.

A similar argument proves the invariance of the modular connectivity numbers P^i (π) and, hence, of the coefficients of torsion.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

GEOMETRIES OF PATHS FOR WHICH THE EQUATIONS OF THE PATHS ADMIT $n(n+1)/2$ INDEPENDENT LINEAR FIRST INTEGRALS*

BY

L. P. EISENHART

1. The paths of a space S_n of coördinates $x^1, \cdots, x^n$ are by definition the integral curves of a system of equations of the form

$$\frac{d^2x^i}{ds^2}+\Gamma^i_{jk}\frac{dx^j}{ds}\frac{dx^k}{ds}=0 \qquad (i,j,k=1,\cdots,n), \tag{1.1}$$

where Γ^i_{jk} are functions of the x's such that $\Gamma^i_{jk}=\Gamma^i_{kj}$, and s is a parameter peculiar to each path. It is understood throughout the paper that a repeated index indicates summation with respect to the index.

If each integral of equations (1.1) satisfies the condition

$$a_i\frac{dx^i}{ds}=\text{const.}, \tag{1.2}$$

where a_i are functions of the x's, equations (1.1) are said to admit a *linear first integral*. A necessary condition is

$$a_{i,j}+a_{j,i}=0, \tag{1.3}$$

where

$$a_{i,j}=\frac{\partial a_i}{\partial x^j}-a_h\Gamma^h_{ij}. \tag{1.4}$$

As thus defined $a_{i,j}$ is a generalized covariant derivative of the covariant vector a_i. It is understood in what follows that a subscript or subscripts preceded by a comma denote generalized covariant derivatives of the first or higher order according to the number of these subscripts. In particular, $\psi_{,i}$ is the derivative $\partial\psi/\partial x^i$.

For this covariant differentiation we have the identities

$$a_{i,jk}-a_{i,kj}=a_h B^h_{ijk}, \tag{1.5}$$

$$a_{i,jkl}-a_{i,jlk}=a_{h,j}B^h_{ikl}+a_{i,h}B^h_{jkl}, \tag{1.6}$$

---

* Presented to the Society, February 27, 1926; received by the editors in December, 1925.

where B^h_{ijk}, defined by

$$B^h_{ijk} = \frac{\partial \Gamma^h_{ik}}{\partial x^j} - \frac{\partial \Gamma^h_{ij}}{\partial x^k} + \Gamma^l_{ik}\Gamma^h_{lj} - \Gamma^l_{ij}\Gamma^h_{lk} \,, \tag{1.7}$$

are the components of a tensor, called the *curvature tensor*.*

From (1.3) we have

$$a_{i,jk} + a_{j,ik} = 0 \,. \tag{1.8}$$

If we add to this equation the analogous equation $a_{k,ij} + a_{i,kj} = 0$ and subtract $a_{j,ki} + a_{k,ji} = 0$, the resulting equation is reducible to

$$a_{i,jk} = -a_l B^l_{kij} \dagger \tag{1.9}$$

by means of (1.5) and the identities

$$B^h_{ijk} + B^h_{ikj} = 0 \,, \tag{1.10}$$

$$B^h_{ijk} + B^h_{jki} + B^h_{kij} = 0 \,, \tag{1.11}$$

which are consequences of (1.7).

When we express the conditions of integrability of equations (1.9) by means of (1.6), we obtain

$$a_h(B^h_{kij,l} - B^h_{lij,k}) + a_{h,p}(\delta^p_l B^h_{kij} - \delta^p_k B^h_{lij} + \delta^p_j B^h_{ikl} - \delta^p_i B^h_{jkl}) = 0 \,, \tag{1.12}$$

where

$$\delta^p_l = 1 \text{ or } 0 \,, \tag{1.13}$$

according as $p = l$ or $p \neq l$.

If (1.2) is to be a first integral, the functions a_i must satisfy (1.3), (1.9) and (1.12), so that in general such a first integral does not exist. By means of (1.9) the second and higher derivatives of the a's are expressible linearly in terms of the a's and their first derivatives. There are $n(n+1)$ of these quantities, and they are subject to the $n(n+1)/2$ conditions (1.3). Hence the solutions of the equations (1.3) and (1.9) involve at most $n(n+1)/2$ arbitrary constants, and this number only in case equations (1.12) are satisfied identically. It is our purpose to determine the character and properties of spaces for which the number of constants is $n(n+1)/2$. In Riemannian

---

* Cf. Eisenhart, Annals of Mathematics, ser. 2, vol. 24 (1923), p. 370.

† Cf. Veblen and Thomas, these Transactions, vol. 25 (1923), p. 592. The change in sign is due to a difference in the definition of B^h_{ijk}.

geometry this is a characteristic property of spaces of constant Riemannian curvature.*

2. Before proceeding to the solution of this problem we observe that by contracting the tensor B^h_{ijl} we obtain

$$B_{ij} = B^h_{ijh} = b_{ij} + \varphi_{ij} , \tag{2.1}$$

$$S_{ij} = B^h_{hij} = -2\varphi_{ij} , \tag{2.2}$$

where b_{ij} and φ_{ij} denote the symmetric and skew-symmetric parts of the tensor B_{ij}, and that φ_{ij} can be shown to be the curl of a vector φ_i, that is,

$$\varphi_{ij} = \frac{\partial\varphi_i}{\partial x^j} - \frac{\partial\varphi_j}{\partial x^i} .\dagger \tag{2.3}$$

In order that equations (1.12) be satisfied identically in consequence of (1.3), it is necessary that

$$B^h_{kij,l} - B^h_{lij,k} = 0 , \tag{2.4}$$

$$\begin{aligned}\delta^p_l B^h_{kij} - \delta^h_l B^p_{kij} - \delta^p_k B^h_{lij} + \delta^h_k B^p_{lij}\\ + \delta^p_j B^h_{ikl} - \delta^h_j B^p_{ikl} - \delta^p_i B^h_{jkl} + \delta^h_i B^p_{jkl} = 0 .\end{aligned} \tag{2.5}$$

Contracting for p and l in equations (2.5), we obtain, in consequence of (1.10), (1.11), (2.1) and (2.2),

$$B^h_{kij} = \frac{1}{n-1}(\delta^h_j B_{ik} - \delta^h_i B_{jk} + 2\delta^h_k \varphi_{ij}) . \tag{2.6}$$

When these equations are contracted for h and j, we get

$$B_{ki} - B_{ik} = \frac{2\varphi_{ik}}{n-1} . \tag{2.7}$$

Comparing this equation with (2.1), we have that $\varphi_{ik}=0$, that is, φ_i in (2.3) is a gradient and the tensor B_{ij} is symmetric. Also (2.6) reduces to

$$B^h_{kij} = \frac{1}{n-1}(\delta^h_j B_{ik} - \delta^h_i B_{jk}) . \tag{2.8}$$

Equations (2.5) are satisfied identically by (2.8), and (2.4) are reducible by (2.8) to

$$B_{ik,l} - B_{il,k} = 0 . \tag{2.9}$$

---

* Cf. Eisenhart, *Riemannian Geometry*, Princeton University Press, 1925, p. 238.

† Cf. Eisenhart, Annals of Mathematics, loc. cit., p. 372.

When the expressions from (2.8) are substituted in the identities*

$$B^h_{kij,l}+B^h_{kjl,i}+B^h_{kli,j}=0$$

we obtain

$$\delta^h_j(B_{ik,l}-B_{lk,i})+\delta^h_l(B_{jk,i}-B_{ik,j})+\delta^h_i(B_{lk,j}-B_{jk,l})=0 \ .$$

Contracting for h and j, we obtain

$$(n-2)(B_{ik,l}-B_{lk,i})=0 \ .$$

Hence when $n\neq 2$ equations (2.8) and that B_{ij} be symmetric are necessary and sufficient conditions of the problem, and equations (2.9) when $n=2$, since in the latter case (2.8) are satisfied identically.

3. If a geometry of paths is defined by a given set of equations (1.1) and we define a set of functions $\bar{\Gamma}^i_{jk}$ by the equations

$$\bar{\Gamma}^i_{jk}=\Gamma^i_{jk}+\delta^i_j\psi_k+\delta^i_k\psi_j \ , \tag{3.1}$$

where ψ_i are the components of an arbitrary vector, and if also we define a parameter $\bar{s}$ along a path as a function of $\bar{s}$ by the equation

$$\frac{d^2s}{d\bar{s}^2}=-2\psi_i\frac{dx^i}{ds}\left(\frac{d\bar{s}}{ds}\right)^2, \tag{3.2}$$

equations (1.1) can be written in the form

$$\frac{d^2x^i}{d\bar{s}^2}+\bar{\Gamma}^i_{jk}\frac{dx^j}{d\bar{s}}\frac{dx^k}{d\bar{s}}=0 \ , \tag{3.3}$$

and this is the most general way in which the $\bar{\Gamma}$'s and $\bar{s}$ can be chosen to give this result.†

If we denote by $\bar{B}^h_{ijk}$ the function of the $\bar{\Gamma}$'s analogous to (1.7), we find that

$$\bar{B}^h_{ijk}=B^h_{ijk}+\delta^h_i(\psi_{k,j}-\psi_{j,k})+\delta^h_k(\psi_{i,j}-\psi_i\psi_j)-\delta^h_j(\psi_{i,k}-\psi_i\psi_k) \ , \tag{3.4}$$

where $\psi_{i,j}$ is the covariant derivative of ψ_i with respect to the Γ's.

Contracting (3.4) for h and k and for h and i, we have

$$\bar{B}_{ij}=B_{ij}+n\psi_{i,j}-\psi_{j,i}-(n-1)\psi_i\psi_j \ , \tag{3.5}$$

* Cf. Veblen and Thomas, loc. cit., p. 580; also, Schouten, *Der Ricci-Kalkül*, Berlin, Springer, 1924, p. 91.

† Weyl, Göttinger Nachrichten, 1921, p. 99; also, Eisenhart, loc. cit., p. 377.

and

$$(3.6)\qquad \begin{aligned}\bar{\varphi}_{jk}&=\varphi_{jk}+\tfrac{1}{2}(n+1)(\psi_{j,k}-\psi_{k,j})\\ &=\varphi_{jk}+\tfrac{1}{2}(n+1)\left(\frac{\partial\psi_j}{\partial x^k}-\frac{\partial\psi_k}{\partial x^j}\right).\end{aligned}$$

The quantities Γ^i_{jk} determine a definition of infinitesimal parallelism in the sense of Levi-Civita and Weyl. Hence the same set of paths lead to different affine connections, according to the choice of the vector ψ_i. Those properties of the space which depend only upon the paths constitute a *projective geometry of paths*, and those depending upon a particular choice of ψ_i an *affine geometry of paths*.

It is readily shown, as was first pointed out by Weyl,* that the tensor

$$(3.7)\qquad W^h_{ijk}=B^h_{ijk}+\frac{1}{n+1}\delta^h_i(B_{jk}-B_{kj})+\frac{1}{n^2-1}[\delta^h_j(nB_{ik}+B_{ki})-\delta^h_k(nB_{ij}+B_{ji})]$$

is independent of the choice of ψ_i. Weyl called it the *projective curvature tensor*.

From (3.6) and (2.3) it is seen that, if we take

$$\psi_i=-\frac{2}{n+1}\left(\varphi_i+\frac{\partial\sigma}{\partial x^i}\right),$$

where σ is any function of the x's, then $\bar{\varphi}_{ij}=0$. Hence we have

The affine connection of a given geometry of paths can be chosen so that the tensor B_{ij} is symmetric.†

When B_{ij} is symmetric, equations (3.7) reduce to

$$(3.8)\qquad W^h_{ijk}=B^h_{ijk}+\frac{1}{n-1}(\delta^h_jB_{ik}-\delta^h_kB_{ij}).$$

Comparing this equation with (2.8), we have the following theorem:

A necessary and sufficient condition that the equations of the paths of a space S_n for $n>2$ admit $n(n+1)/2$ independent linear first integrals is that the tensor W^h_{ijk} vanish and the tensor B_{ij} be symmetric.

4. Weyl has shown‡ that for $n>2$ the vanishing of the tensor W^h_{ijk} is a necessary and sufficient condition that a vector ψ_i can be chosen so that

* Loc. cit., p. 101.

† Cf. Eisenhart, loc. cit., p. 378.

‡ Loc. cit., pp. 103, 105.

for the new affine connection the curvature tensor $\bar{B}^h_{ijk}$ is a zero tensor and that the vanishing of the latter tensor is a necessary and sufficient condition that a coördinate system exist, which we call *cartesian*, for which all of the $\bar{\Gamma}$'s are zero. He has called a space satisfying the former conditions *projective plane*.

Weyl has shown* also that when $n=2$ the tensor W^h_{ijk} vanishes identically and that equations (2.9) are necessary and sufficient conditions that the space be projective plane. Hence we have

A necessary and sufficient condition that the equations of the paths of any space admit $n(n+1)/2$ independent linear first integrals is that the space be projective plane and that the tensor B_{ij} be symmetric.

From the results of Weyl it follows that a vector ψ_i can be chosen so that $\bar{B}_{ij}$ is a zero tensor, and from (3.5) that this vector is a gradient, if B_{ij} is symmetric. In the coördinate system for which the $\bar{\Gamma}$'s are zero, we have from (3.1)

$$\Gamma^i_{jk} = -(\delta^i_j \psi_{,k} + \delta^i_k \psi_{,j}) \; . \tag{4.1}$$

Conversely, when we take the Γ's in the form (4.1), where ψ is an arbitrary function, we have

$$B^h_{ijk} = e^{-\psi}\left(\delta^h_j \frac{\partial^2 e^{\psi}}{\partial x^i \partial x^k} - \delta^h_k \frac{\partial^2 e^{\psi}}{\partial x^i \partial x^j}\right) . \tag{4.2}$$

Contracting for h and k, we have

$$B_{ij} = (1-n)e^{-\psi}\frac{\partial^2 e^{\psi}}{\partial x^i \partial x^j} , \tag{4.3}$$

from which it follows that the conditions (2.8) are satisfied.

For the expressions (4.1) of the Γ's we have

$$B_{ij,k} - B_{ik,j} = \frac{\partial B_{ij}}{\partial x^k} - \frac{\partial B_{ik}}{\partial x^j} + B_{ij}\psi_{,k} - B_{ik}\psi_{,j} \; . \tag{4.4}$$

When the expressions (4.3) are substituted, we find that (2.9) are satisfied. Hence we have

The most general geometries of paths for which the equations of the paths admit $n(n+1)/2$ linear first integrals are defined by (4.1) *in which $\psi_{,i}$ is the gradient of an arbitrary function ψ.*

---

* Loc. cit., p. 104.

5. In the coördinate system for which the Γ's have the form (4.1) equations (1.3) become

$$\frac{\partial b_i}{\partial x^j} + \frac{\partial b_j}{\partial x^i} = 0 , \tag{5.1}$$

where

$$b_i = a_i e^{2\psi} . \tag{5.2}$$

Equations (5.1) are the form which (1.3) assume in a euclidean space referred to cartesian coördinates. In this case equations (1.9) become

$$\frac{\partial^2 b_i}{\partial x^j \partial x^k} = 0 . \tag{5.3}$$

From (5.1) for $j = i$ it follows that b_i is independent of x^i, and from (5.1) and (5.3) that the general solution is

$$b_i = c_{ij} x^j + d_i , \tag{5.4}$$

where c_{ij} and d_i are arbitrary constants, subject to the condition that c_{ij} is skew-symmetric in the indices. Hence there are $n(n+1)/2$ arbitrary constants as desired, and the a's are given by (5.2) and (5.4).

6. Let S_n be a space for which B_{ij} is symmetric and also conditions (2.8) and (2.9) are satisfied; and consider the equations

$$\frac{\partial^2 \theta}{\partial x^i \partial x^j} - \Gamma^k_{ij} \frac{\partial \theta}{\partial x^k} = \frac{1}{n-1} B_{ij} \theta , \tag{6.1}$$

which may be written

$$\theta_{,ij} = \frac{1}{n-1} B_{ij} \theta . \tag{6.2}$$

The conditions of integrability of these equations, namely

$$\theta_{,ijk} - \theta_{,ikj} = \theta_{,h} B^h_{ijk} ,$$

are reducible by (6.2) to

$$\theta_{,h} \left[B^h_{ijk} - \frac{1}{n-1} \left(\delta^h_k B_{ij} - \delta^h_j B_{ik} \right) \right] - \frac{\theta}{n-1} \left(B_{ij,k} - B_{ik,j} \right) = 0 .$$

Since these conditions are satisfied identically in consequence of (2.8) and (2.9), equations (6.2) are completely integrable and a solution is determined by arbitrary values of the $n+1$ quantities θ and $\theta_{,i}$ for initial values of the x's, that is, the complete solution of (6.2) involves $n+1$ arbitrary constants.

Consequently $n+1$ solutions $\varphi^\alpha(x^1, \cdots, x^n)$ for $\alpha=1, \cdots, n+1$ of equations (6.2) exist for which the determinant

$$(6.3)\qquad \begin{vmatrix} \dfrac{\partial\varphi^1}{\partial x^1} & \cdots & \dfrac{\partial\varphi^1}{\partial x^n} & \varphi^1 \\ \dfrac{\partial\varphi^2}{\partial x^1} & \cdots & \dfrac{\partial\varphi^2}{\partial x^n} & \varphi^2 \\ \cdot & \cdots & \cdot & \cdot \\ \dfrac{\partial\varphi^{n+1}}{\partial x^1} & \cdots & \dfrac{\partial\varphi^{n+1}}{\partial x^n} & \varphi^{n+1} \end{vmatrix}$$

is different from zero and the matrix of the first n columns is of rank n. Hence the jacobian of the equations

$$(6.4)\qquad y^\alpha = x^{n+1}\varphi^\alpha(x^1, \cdots, x^n) \qquad (\alpha=1, \cdots, n+1)$$

is different from zero, and these equations define a transformation of coördinates in a space S_{n+1}.

We define an affine connection in the S_{n+1} in coördinates x^α by taking for $\Gamma^i_{jk}(i, j, k=1, \cdots, n)$ the expressions for these functions for the given S_n and in addition

$$(6.5)\qquad \Gamma^{n+1}_{ij} = \frac{1}{n-1}B_{ij}x^{n+1}, \quad \Gamma^\alpha_{n+1\,i} = \frac{\delta^\alpha_i}{x^{n+1}}, \quad \Gamma^\alpha_{n+1\;n+1} = 0 \begin{pmatrix} i, j=1, \cdots, n\,; \\ \alpha=1, \cdots, n+1 \end{pmatrix}.$$

If $\overline{\Gamma}^\alpha_{\beta\gamma}$ denote the coefficients of the affine connection in the y's we have

$$\frac{\partial^2 y^\alpha}{\partial x^\beta \partial x^\gamma} + \overline{\Gamma}^\alpha_{\mu\nu}\frac{\partial y^\mu}{\partial x^\beta}\frac{\partial y^\nu}{\partial x^\gamma} = \Gamma^\mu_{\beta\gamma}\frac{\partial y^\alpha}{\partial x^\mu} \qquad (\alpha, \beta, \mu, \nu=1, \cdots, n+1).$$

From these equations we have in consequence of (6.1), (6.4) and (6.5)

$$(6.6)\qquad \overline{\Gamma}^\alpha_{\mu\nu}\frac{\partial y^\mu}{\partial x^\beta}\frac{\partial y^\nu}{\partial x^\gamma} = 0\,.$$

Since the jacobian of the transformation (6.4) is different from zero, equations (6.6) are equivalent to $\overline{\Gamma}^\alpha_{\mu\nu}=0$. Consequently S_{n+1} as defined is a euclidean or flat space and the y's are cartesian coördinates.

From the definition of the affine connection of S_{n+1} it follows that the affine connection induced in the hypersurface $x^{n+1}=1$, that is,

$$(6.7)\qquad y^\alpha = (\varphi^\alpha x,^1 \cdots, x^n)$$

is that of the given S_n. Consequently we have

A space S_n whose equations of the paths admit $n(n+1)/2$ independent linear first integrals is a hypersurface of a flat-space of $n+1$ dimensions.

When the coördinates x^i in S_n are such that the Γ's have the form (4.1), equations (6.1) are reducible in consequence of (4.3) to

$$\frac{\partial^2(\theta e^{\psi})}{\partial x^i \partial x^j} = 0 .$$

Hence the equations (6.7) in this coördinate system are

$$y^{\alpha} = e^{-\psi}(\overset{\alpha}{a}_i x^i + b^{\alpha}) , \tag{6.8}$$

where the a's and b's are arbitrary constants subject to the condition that the rank of the jacobian matrix $||\partial y^{\alpha}/\partial x^i||$ is n and the determinant (6.3) is different from zero.

PRINCETON UNIVERSITY,
PRINCETON, N. J.

MULTIPLY TRANSITIVE SUBSTITUTION GROUPS*

BY

G. A. MILLER

1. Introduction

A substitution group G of degree n is said to be r-fold transitive if each of the $n(n-1)\cdots(n-r+1)$ permutations of its n letters taken r at a time is represented by at least one substitution of G. It is not sufficient to say that each of its possible sets of r letters is transformed into every one of these sets by the substitutions of G. For instance, in each of the semi-metacyclic groups of degree p, p being a prime number of the form $4n+3$, every possible pair of its letters is transformed into every other such pair by the substitutions of the group and yet these groups are only simply transitive. Another well known definition of an r-fold transitive group of degree n is that this group contains a transitive subgroup of each of the degrees n, $n-1, \cdots, n-r+1$. It is obvious that these two definitions are equivalent and that if a group is r-fold transitive it is also $(r-\alpha)$-fold transitive, where α is a positive integer $\leqq r-1$.

It is not difficult to prove that a necessary and sufficient condition that a primitive group of degree n be r-fold transitive, $r>1$, is that it contain a doubly transitive group of degree $n-r+2$. This may have to coincide with the group itself if the group is only doubly transitive. To prove that such a group G must contain a triply transitive group of degree $n-r+3$, whenever $r>2$ it is only necessary to note that G involves a substitution which transforms at least one letter of the doubly transitive group of degree $n-r+2$ into a letter of this group while it transforms another letter of this group into a letter not involved therein since G is primitive. Hence G involves at least two primitive subgroups of degree $n-r+2$ which have at least one common letter but do not have all their letters in common. It must therefore involve two such primitive groups which have all except one letter of each in common. These two groups obviously generate a triply transitive group since they themselves are doubly transitive. When $r>3$ we may repeat this argument and prove the existence in G of a four-fold transitive group of degree $n-r+4$, etc. It must therefore contain a transitive group of each of the degrees $n, n-1, \cdots, n-r+1$.

---

* Presented to the Society, September 11, 1925; received by the editors in December, 1925.

The number of the largest transitive subgroups of degree $n-\alpha$, $\alpha \leqq r-1$, is obviously $n(n-1) \cdots (n-\alpha+1)$ and each of these conjugate subgroups is invariant under an intransitive subgroup of degree n which has for one of its transitive constituents the symmetric group of degree α. A necessary and sufficient condition that G be symmetric is that this intransitive group be the direct product of its transitive constituents, and if one such intransitive group is the direct product of its transitive constituents every one of them must have this property. In all other multiply transitive groups these intransitive groups, for $\alpha=2$, have at least one transitive constituent whose order is equal to the order of this intransitive group. A necessary and sufficient condition that all these intransitive groups are dimidiations, when $n>4$, is that G is an alternating group. In this case they are always dimidiations between symmetric groups. Hence the following theorem:

If G is an r-fold transitive group of degree n then each of its subgroups of order g_α and of degree $n-\alpha$, $\alpha \leqq r-1$, which is not contained in a larger subgroup of this degree, is transitive and is invariant under an intransitive group of degree n and of order $g_\alpha \cdot \alpha!$ but under no larger group. This intransitive group has for one of its transitive constituents the symmetric group of degree α, and its other transitive constituent is simply isomorphic with this intransitive group whenever G is not symmetric or alternating, and only then when $\alpha>2$.

From the preceding theorem it results directly that whenever $r>3$ the largest transitive groups of degree $n-r+1$ contained in G cannot be cyclic since the group of isopmorphisms of a cyclic group is abelian. In fact when G is not alternating or symmetric, the largest transitive subgroups of degree $n-\alpha$ must be invariant under a group of this degree whose order is $\alpha!$ times the order of this largest transitive group, and the corresponding quotient group must be the symmetric group of degree α. In particular, when $n-\alpha$ is a prime number p then $n \leqq p+2$. When G is exactly r-fold transitive, $n-2>r>1$, it can obviously not contain a transitive group of degree $n-r$ but it may possibly contain such groups of a lower degree. If such a transitive group appears in G it must be imprimitive and G must involve also imprimitive groups of a larger degree. In fact, the largest transitive subgroups of the former degree are known to be imprimitive, and since G is primitive two such subgroups which have the largest possible number of common letters must have a common system of imprimitivity since they cannot generate a group whose degree is only one larger than their common degree. The letters which appear in one of these subgroups but not in the other form one set of a system of imprimitivity of this group.

Whenever a substitution group of degree n transforms every possible set of r, $r<n$, of its letters into every other such set the group must be primitive unless $r=1$ or $n-1$. Since a transitive group of degree n could be defined as a group in which every possible set of $n-1$ of its letters is transformed into every other such set, it results directly that an r-fold transitive group could not be defined as one in which all the possible sets composed of r of its letters are transformed transitively under the group. The fact that when $1<r<n-1$ every group which transforms all its possible sets of r letters transitively must be primitive results almost directly from the definitions of intransitive and primitive groups, since in the latter groups it would be possible to choose two such sets in such a manner that they would obviously not be conjugate under the group. Hence the following theorem:

If a substitution group of degree n has the property that every possible set of r of its letters, $1<r<n-1$, *is transformed transitively under the group, then this group must be primitive.*

2. Regular subgroups of odd prime order

It is well known that a primitive group of degree n which is not alternating or symmetric cannot contain a regular subgroup of odd prime degree p unless n has one of the three values p, $p+1$, $p+2$. In particular, such a primitive group G of degree $p+k$ cannot be more than k times transitive whenever $k>2$. We proceed to establish the following theorem:

If a primitive group involves a regular subgroup of odd prime order p, all such subgroups generate a simple group unless p is the form $2^{\alpha}-1$. *In this case they generate either a simple group of composite order or a subgroup of the holomorph of the abelian group of order* 2^{α} *and of type* $(1, 1, 1, \cdots)$.

As regards the alternating and the symmetric groups this theorem is obviously true. In fact, in this case these regular subgroups always generate the alternating group whose degree is equal to the degree of the group. When the degree of G is p the theorem is well known and obvious. When the degree of G is $p+1$ it must be multiply transitive and to an invariant subgroup of G there must correspond an invariant subgroup of the group composed of all the substitutions of G which omit one letter. Since the latter is generated by operators of order p whenever the former is thus generated it must be simple and the former contains no invariant subgroup except the identity and possibly one of order $p+1$. The latter can exist only when $p+1=2$, since the only regular group of order $p+1$ which has an operator of order p in its group of isomorphisms is the abelian group of order 2^{α} and of the type $(1, 1, 1, \cdots)$.

It remains to consider the case when the degree of G is $p+2$. Since the regular subgroups of order p are transformed into themselves by twice as many substitutions under G as under a subgroup composed of all the substitutions of G which omit a given letter, it results that $p+1$ cannot be divisible by 4. That is, *a triply transitive group of degree $p+2$, p being an odd prime number, must be composed of positive substitutions, and such a group cannot exist unless p is of the form $4k+3$.* If G exists it cannot involve an invariant subgroup of order $p+2$ since a group of this order cannot have an operator of order p in its group of isomorphisms. Hence G cannot contain an invariant subgroup of order $(p+1)$ $(p+2)$. That is, the operators of order p contained in G generate a simple group of composite order and the quotient group of G with respect to this invariant subgroup is a cyclic group of odd order. This order is a divisor of $p-1$.

A triply transitive group of degree $p+2$ is completely determined by any one of its maximal doubly transitive subgroups of degree $p+1$ in the sense that only one triply transitive group of degree $p+2$ can involve a given group of degree $p+1$ as the subgroup composed of all its substitutions which omit a given letter. In particular, the number of the triply transitive substitution groups of degree $p+2$ can certainly not exceed the number of the doubly transitive groups of degree $p+1$. It may also be noted that if the doubly transitive groups of degree $p+1$ are given it is very easy to determine all the possible triply transitive groups of the degree $p+2$, since it is only necessary to take any one of the p substitutions of order 2 and degree $p-1$ which transform into its inverse an arbitrary substitution of order p contained in this doubly transitive group and annex to this substitution a transposition composed of the remaining letter of this doubly transitive group and an arbitrary letter not found therein. A necessary and sufficient condition that the triply transitive group of degree $p+2$ exist is that the given substitution and the given doubly transitive group generate a group whose order does not exceed $p+2$ times the order of this doubly transitive group.

Just as a characteristic subgroup which appears in every other characteristic subgroup but is not the identity has been called the fundamental characteristic subgroup so we may define the term *fundamental invariant subgroup* as an invariant subgroup which is not the identity but appears in every other possible invariant subgroup of the group. From the preceding developments it follows almost directly that *if a substitution composed of a single cycle involving an odd prime number p of letters appears in a primitive group then all such substitutions found in this group generate a fundamental invariant subgroup of the group unless $p+1$ is of the form 2^α and the group is*

contained in the holomorph of the abelian group of order 2^a *and of type* (1, 1, 1, $\cdots$).

3. Substitutions composed of two cycles of odd prime order

If a primitive group G which is neither alternating nor symmetric involves a substitution s_1 *composed of two cycles of prime order p the degree of G cannot exceed* $2p+6$. Since the primitive groups whose class does not exceed six are all well known,* we shall assume in what follows that $p>3$. Suppose that the degree of G is larger than $2p+6$. Since G is primitive it contains a substitution s_2 which is similar to s_1 and involves at least one letter of s_1 and also at least one letter not found in s_1. The group generated by s_1 and s_2 cannot have three systems of intransitivity since it cannot contain a transitive constituent of order p, nor can it contain a substitution composed of a single cycle of order p according to the preceding section. The proof of the fact that it cannot contain a transitive constituent of order p results almost directly from the following theorem: *If a group contains two transitive subgroups which have at least one letter in common it must also contain two conjugate transitive subgroups which have all their letters in common except possibly those found in a system of imprimitivity of one of these groups. In particular, if one of these groups is primitive the given transitive subgroups involve all except possibly one of the letters of this primitive group.*

If the group generated by s_1 and s_2 is transitive it may be assumed according to the theorem stated above that the degree of this group is either $2p+1$ or $2p+2$. In the former case it would have to be primitive, and hence this is impossible. If in the latter case it is imprimitive its systems of imprimitivity are permuted under the group according to a doubly transitive group of degree $p+1$. If this group of degree $2p+2$ could be found in an imprimitive group of degree $2p+4$, its systems of imprimitivity would be transformed according to a triply transitive group of degree $p+2$. According to the preceding section this imprimitive group of degree $2p+4$ could not appear in an imprimitive group of degree $2p+6$ and hence the group of degree $2p+6$ would be at least doubly transitive. Since the substitution s_1 can not appear in a group of degree $2p+k$, $k>2$, which is as much as k times transitive, it has been proved that the group generated by s_1 and s_2 cannot be assumed to be transitive. It should be noted that the proof of the theorem that a group of degree $2p+k$, $k>2$, cannot be more than k times transitive implies that such a group cannot be as much as k times transitive when it involves a substitution composed of two cycles of order p.†

* Cf. W. A. Manning, these Transactions, vol. 4 (1903), p. 351.

† G. A. Miller, Bulletin of the American Mathematical Society, vol. 4 (1898), p. 143.

It remains to consider the case when every pair of substitutions such as s_1 and s_2 generates a group which has two and only two transitive constituents and the order of each transitive constituent exceeds p. By extending the group generated by s_1 and s_2 by the other similar substitutions, we cannot construct a group whose degree exceeds $2p+4$ without arriving at an intransitive group which has alternating transitive constituents. Such an intransitive group would have a class which could not exceed 6 and must therefore be excluded. Hence we have the result, just as in the preceding cases, that the degree of G cannot exceed $2p+6$, and the theorem announced in the first paragraph of this section has been established.

4. Substitutions composed of three or more cycles of odd prime order

Suppose that G is of degree $3p+k$, $p>3$, $k>3$, and assume that G is k-fold transitive. From the preceding section it results that there is no substitution in G composed of two cycles of order p, and hence the order of G is not divisible by p^2 when $k>5$. If G involves a substitution s composed of three cycles of order p, all of its subgroups of order p must be conjugate under G. In fact, the subgroups of order p which are contained in a group H composed of all the substitutions of G which omit k letters must be conjugate under H. Hence the subgroup composed of all the substitutions of G which transform into itself a subgroup of order p must have for one of its transitive constituents the symmetric group of degree k. To the identity of this constituent there must correspond the subgroup of H which transforms into itself the given subgroup of order p.

As k was assumed to be 4, for the present, it is obvious that to the invariant subgroup of order 4 in the symmetric constituent of degree k there has to correspond a subgroup involving substitutions in each co-set corresponding to this subgroup of order 4 which do not interchange the systems of intransitivity of the given subgroup of order p. Since this subgroup of order 4 is non-cyclic while the group of isomorphisms of the group of order p is cyclic, it results that G must involve substitutions containing no more than four letters. Since this is impossible, it has been proved that G cannot exist for $k=4$ and hence it cannot exist for larger values of k. It has therefore been proved that a substitution group of degree $3p+k$, $k>3$, which is k-fold transitive cannot involve any substitution composed of three cycles of odd prime order. *In particular, a group of degree $3p+k$, $p>3$, $k>3$, cannot be more than k-fold transitive.*

Suppose that G is of degree $lp+k$, $p>3$, $p>k>l$, and suppose that G contains a substitution of order p and of degree lp. If G is k-fold transitive

each of its Sylow subgroups of order p^m is invariant under an intransitive group which has for one of its transitive constituents the symmetric group of degree k. Each one of the l systems of intransitivity of the Sylow subgroup of order p^m is transformed into itself by at least some of the substitutions which correspond to each one of the substitutions of the alternating group of degree k contained in the given symmetric constituents of degree k. Hence it results that G must involve the substitutions of this alternating group. This is obviously impossible and hence G cannot be k-fold transitive if it involves substitutions of order p and of degree lp. If G were $(k+1)$-fold transitive it would clearly involve such substitutions. We have therefore established the following theorem:

*If a substitution group is of degree $lp+k$, p being a prime number, $p>l<k$, then this group cannot be more than k-fold transitive, $k>2$, unless it is the alternating or the symmetric group.**

If in the preceding theorem we assume that $p>\sqrt{n}$, where n is the degree of G, it results that k may be selected so as not to exceed $p+l+1$. In fact, if k were greater than this number, we could increase the value of l by unity and thus reduce the value of k by p, provided l is less than $p-1$, as it will be when p is properly chosen, according to the well known theorem that for every integer $x>7$ there is at least one prime p such that

$$x/2<p\leqq x-3\ .†$$

Hence it results that when p is a prime number which exceeds the smallest integer greater than $\sqrt{n}$ it may be assumed that p is not larger than $2\sqrt{n}-1$. Hence k can always be so selected as not to exceed $5\sqrt{n}/2$. For large values of n this obviously gives a much smaller upper limit for the multiplicity of a transitive group than the limit commonly given, viz., $n/3+1$.‡

It should be added that in view of the fact that the prime numbers are usually much closer together than the given formula due to P. L. Tschebychef indicates, the theorem given at the close of the preceding paragraph is much more useful than the formula $n/3+1$ for the determination of the upper limit of the multiplicity of transitivity. For instance, p may be so selected that by means of this theorem it follows directly that no group whose degree does not exceed 200 can be more than 8-fold transitive without involving the alternating group.

* G. A. Miller, Bulletin of the American Mathematical Society, vol. 22 (1915), p. 70.
† Cf. G. A. Miller, School Science and Mathematics, vol. 21 (1921), p. 874.
‡ Cf. Pascal's *Repertorium der höheren Mathematik*, vol. 1, 1910, p. 211.

University of Illinois,
Urbana, Ill.

AN EXTENSION OF LAGRANGE'S EXPANSION*

BY

H. BATEMAN

1. **Solution of a partial differential equation by means of a definite integral.** Soon after the publication of Whittaker's solution of Laplace's equation† it was realized that the linear homogeneous partial differential equation

$$f\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)V=0 \tag{1}$$

can be satisfied by a definite integral of type

$$V=\int_0^\theta F[\sigma,\tau]d\tau\ ,$$

where

$$\sigma=x\xi(\tau)+y\eta(\tau)+z\zeta(\tau)-\chi(\tau)$$

by simply choosing functions $\xi(\tau)$, $\eta(\tau)$, $\zeta(\tau)$ such that‡

$$f[\xi(\tau)\ ,\eta(\tau)\ ,\zeta(\tau)]=0$$

and making the upper limit θ a constant.

It has been known for some time§ that if the upper limit θ is not a constant but a function of x, y, and z defined by the equation

$$x\xi(\theta)+y\eta(\theta)+x\zeta(\theta)=\chi(\theta) \tag{2}$$

the differential equation still may be satisfied. A direct verification of this result for the case when $f(\xi,\eta,\zeta)$ is any homogeneous polynomial of degree m in ξ, η and ζ is rather tedious but it may be accomplished by establishing the following rule for the differentiation of V.

---

* Presented to the Society, San Francisco Section, April 4, 1925; received by the editors in April, 1925.

† Monthly Notices, Royal Astronomical Society, vol. 62 (1902); Mathematische Annalen, 1903, p. 333.

‡ H. Bateman, Proceedings of the London Mathematical Society, ser. 2, vol. 1 (1904), p. 451.

§ A particular case is mentioned in the author's *Electrical and Optical Wave Motion*, p. 12.

Let us write

$$F_n[\sigma,\tau]=\frac{\partial^n}{\partial\sigma^n}F[\sigma,\tau]\ ,$$

$$M=\chi'(\theta)-x\xi'(\theta)-y\eta'(\theta)-z\zeta'(\theta)\ ,$$

$$\Phi(\tau)=f[\xi(\tau)\ ,\ \eta(\tau)\ ,\ \zeta(\tau)]\ ,$$

where at present f is any homogeneous polynomial of degree m; then

$$\begin{aligned} f\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)V= & \int_0^\theta F_m[\sigma,\tau]\Phi(\tau)d\tau+\frac{1}{M}F_{m-1}[0,\theta]\Phi(\theta) \\ & +\frac{1}{M}\frac{\partial}{\partial\theta}\left\{\frac{1}{M}F_{m-2}[0,\theta]\Phi(\theta)\right\}+\cdots \\ & +\left(\frac{1}{M}\frac{\partial}{\partial\theta}\right)^{m-1}\left\{\frac{1}{M}F[0,\theta]\Phi(\theta)\right\}. \end{aligned} \tag{3}$$

This formula clearly shows that if $\Phi(\tau)\equiv 0$ the expression for V satisfies the differential equation (1).

The formula indicates that if

$$W=\int_0^\omega F[\rho,\tau]d\tau\ ,$$

where

$$\rho=(x+a)\xi(\tau)+(y+b)\eta(\tau)+(z+c)\zeta(\tau)-\chi(\tau)\ ,$$

$$(x+a)\xi(\omega)+(y+b)\eta(\omega)+(z+c)\zeta(\omega)=\chi(\omega)\ ,$$

the Taylor expansion of W in powers of a, b and c is

$$\begin{aligned} W= & \int_0^\theta F[\sigma,\tau]d\tau+\int_0^\theta F_1[\sigma,\tau]\phi(\tau)d\tau+\frac{1}{M}F[0,\tau]\phi(\theta) \\ & +\frac{1}{2!}\int_0^\theta F_2[\sigma,\tau][\phi(\tau)]^2d\tau+\frac{1}{2!}[\phi(\theta)]^2\frac{1}{M}F_1[0,\theta] \\ & +\frac{1}{2!}\frac{1}{M}\frac{\partial}{\partial\theta}\left\{\frac{1}{M}[\phi(\theta)]^2F[0,\theta]\right\} \\ & +\frac{1}{3!}\int_0^\theta F_3[\sigma,\tau][\phi(\tau)]^3d\tau+\frac{1}{3!}[\phi(\theta)]^3\frac{1}{M}F_2[0,\theta] \\ & +\frac{1}{3!}\frac{1}{M}\frac{\partial}{\partial\theta}\left\{\frac{1}{M}[\phi(\theta)]^3F_1(0,\theta)\right\} \\ & +\frac{1}{3!}\frac{1}{M}\frac{\partial}{\partial\theta}\left\{\frac{1}{M}\frac{\partial}{\partial\theta}\left(\frac{1}{M}[\phi(\theta)]^3F(0,\theta]\right)\right\} \\ & +\cdots\ , \end{aligned} \tag{4}$$

where

$$\phi(\tau)=a\xi(\tau)+b\eta(\tau)+c\zeta(\tau) .$$

The form of this expansion reminds one of Lagrange's well known expansion and suggests the following theorem.*

2. **Extension of Lagrange's expansion.** *Let z be defined by the equation*

$$z=a+x\phi(z) ;$$

then under suitable conditions we have the expansion

$$\begin{aligned}
\int_0^z F[a+x\phi(\tau)-\tau,\tau]d\tau &= \int_0^a F[a-\tau,\tau]\,d\tau \\
&+\frac{x}{1!}\Big[\int_0^a F_1[a-\tau,\tau]\phi(\tau)d\tau+F[0,a]\phi(a)\Big] \\
&+\frac{x^2}{2!}\Big[\int_0^a F_2[a-\tau,\tau][\phi(\tau)]^2d\tau+F_1[0,a][\phi(a)]^2 \\
&\qquad+\frac{d}{da}\{F[0,a][\phi(a)]^2\}\Big] \qquad (1)\\
&+\frac{x^3}{3!}\Big[\int_0^a F_3[a-\tau,\tau][\phi(\tau)]^3d\tau+F_2[0,a][\phi(a)]^3 \\
&\qquad+\frac{d}{da}\{F_1[0,a][\phi(a)]^3\}+\frac{d^2}{da^2}\{F[0,a][\phi(a)]^3\}\Big] \\
&+\cdots ,
\end{aligned}$$

where as before

$$F_n[\sigma,\tau]=\frac{\partial^n}{\partial\sigma^n}F[\sigma,\tau] .$$

To investigate this expansion we shall endeavor to represent the definite integral on the left hand side by a contour integral

$$I=\frac{1}{2\pi i}\int_C\frac{1-x\phi'(s)}{s-x\phi(s)-a}\,ds\int_0^s F[a+x\phi(\tau)-\tau,\tau]d\tau ,$$

where C is a simple contour enclosing the point $s=a$ and also just one root of the equation

$$s-x\phi(s)=a . \qquad (2)$$

---

* The method used for this simplified form of the theorem can readily be extended so as to be applicable to the general expansion. In §4 the extension is given in detail for the case in which $F[\sigma,\tau]=1$.

When the contour C is given we may choose x so that for all points on the perimeter of C we have the inequality

$$|x\phi(s)|<|s-a| ;$$

then, just as in Hermite's proof of Lagrange's theorem,* it can be shown that the equation (2) has just one root within C.

Let us now suppose that the functions $\phi(s)$ and $F[a+x\phi(s)-s,\, s]$ are analytic within C and on its boundary for values of x that enter into consideration; then expanding by Taylor's theorem we have

$$\begin{aligned}F[a+x\phi(\tau)-\tau,\tau]&=F[a-\tau,\tau]+\frac{x}{1!}\phi(\tau)F_1[a-\tau,\tau]\\&+\frac{x^2}{2!}[\phi(\tau)]^2F_2[a-\tau,\tau]+\cdots+\frac{x^n}{n!}[\phi(\tau)]^nF_n[a-\tau,\tau]\\&+\frac{1}{n!}\int_a^{a+x\phi(\tau)}[x\phi(u)+a-u]^nF_{n+1}[a-u,\tau]du\ ,\end{aligned}$$

$$\begin{aligned}\frac{1-x\phi'(s)}{s-x\phi(s)-a}&=\frac{1}{s-a}-x\frac{d}{ds}\frac{\phi(s)}{s-a}-\frac{1}{2}x^2\frac{d}{ds}\left[\frac{\phi(s)}{s-a}\right]^2-\cdots\\&-\frac{1}{n}x^n\frac{d}{ds}\left[\frac{\phi(s)}{s-a}\right]^n+\frac{x^{n+1}}{(s-a)^{n+1}}\left[\frac{\{\phi(s)\}^{n+1}}{s-a-x\phi(s)}\right.\\&\left.+\phi'(s)\{\phi(s)\}^n\right] .\end{aligned}$$

The coefficient of x^n in the Taylor expansion of I in powers of x is thus

$$\begin{aligned}\frac{1}{2\pi i}\int_C\Big\{&\frac{ds}{s-a}\int_0^s\frac{1}{n!}[\phi(\tau)]^nF_n[a-\tau,\tau]d\tau\\&-ds\frac{d}{ds}\left[\frac{\phi(s)}{s-a}\right]\int_0^s\frac{d\tau}{(n-1)!}\,[(\phi\tau)]^{n-1}F^{n-1}[a-\tau,\tau]\\&-\frac{1}{2}\,ds\frac{d}{ds}\left[\frac{\phi(s)}{s-a}\right]^2\int_0^s\frac{d\tau}{(n-2)!}[\phi(\tau)]^{n-2}F_{n-2}[a-\tau,\tau]-\cdots\\&-\frac{1}{n}\,ds\frac{d}{ds}\left[\frac{\phi(s)}{s-a}\right]^n\int_0^s d\tau F[a-\tau,\tau]\Big\} .\end{aligned}$$

The first term in this series may be evaluated at once by Cauchy's theorem while the other terms may be evaluated by Cauchy's theorem after they

---

* Hedrick-Dunkel, Goursat's *Mathematical Analysis*, vol. 2, Part 1, p. 106.

have been integrated by parts, the terms that are complete differentials disappearing in an integration round a closed contour. The result is

$$\frac{1}{n!}\Bigg[\int_0^a [\phi(\tau)]^n F_n[a-\tau,\tau]d\tau + \binom{n}{1}[\phi(a)]^n F_{n-1}[0,a]$$
$$+\binom{n}{2}\left[\frac{d}{ds}\{[\phi(s)]^n F_{n-2}[a-s,s]\}\right]_{s=a} + \cdots$$
$$+\binom{n}{n}\left[\frac{d^{n-1}}{ds^{n-1}}\{[\phi(s)]^n F[a-s,s]\}\right]_{s=a},$$

where

$$\binom{n}{m} = \frac{n!}{m!(n-m)!}.$$

Let us now write

$$\frac{d}{ds} = D_1 + D_2 ,$$

where D_1 operates only on the s in the first argument of a function $F_k(a-s,s)$, while D_2 operates on the s in $\phi(s)$ and also on the s in the second argument of $F_k(a-s,s)$. The coefficient of

$$\frac{1}{n!}\frac{d^k}{da^k}\{F_{n-k-1}[0,a][\phi(a)^n]\}$$

in the above expression is then found to be

$$C_n = \binom{n}{k+1}\binom{k}{k} - \binom{n}{k+2}\binom{k+1}{k} + \cdots + (-1)^{n-k-1}\binom{n}{n}\binom{n-1}{k}.$$

Making use of the identity

$$\binom{n+1}{r} = \binom{n}{r} + \binom{n}{r-1}$$

we find that

$$C_{n+1} = C_n + \binom{n}{k}(1-1)^{n-k}$$
$$= C_n , \qquad n > k .$$

But

$$C_{k+1} = 1 .$$

Therefore

$$C_n = 1, \qquad n > k .$$

The term involving x^n is thus

$$\frac{x^n}{n!}\Bigg[\int_0^a F_n[a-\tau,\tau][\phi(\tau)]^n d\tau + F_{n-1}[0,a][\phi(a)]^n$$

$$(3)\qquad + \frac{d}{da}\{F_{n-2}[0,a][\phi(a)]^n\} + \cdots + \frac{d^{n-1}}{da^{n-1}}\{F[0,a][\phi(a)[^n\}\Bigg],$$

and so the expansion has the form given above.

It is easy to obtain a formula for the remainder after the term (3) but the formula is complicated and hardly worth writing down. An expression for the remainder may be obtained also by remarking that the integral

$$\int_0^z F[a+x\phi(\tau)-\tau,\tau]d\tau$$

is a function of z and, when this function is analytic, Lagrange's theorem may be applied to it in the usual way and a formula for the remainder written down.

The following alternative method of obtaining the coefficients in the expansion has been developed from a suggestion made by an editor of this journal.

Differentiating the expression

$$I = \int_0^z F[a+x\phi(\tau)-\tau,\tau]d\tau$$

n times with respect to x we obtain

$$\frac{d^nI}{dx^n} = \frac{d^{n-1}}{dx^{n-1}}\left[\frac{dz}{dx}F(0,z)\right] + \frac{d^{n-2}}{dx^{n-2}}\left[\frac{dz}{dx}\phi(z)F_1(0,z)\right]$$

$$+ \frac{d^{n-3}}{dx^{n-3}}\left[\frac{dz}{dx}\{\phi(z)\}^2F_2(0,z)\right] + \cdots$$

$$+ \int_0^z F_n[a+x\phi(\tau)-\tau,\tau]\{\phi(\tau)\}^n d\tau .$$

Putting $x=0$ to calculate the coefficients in the Maclaurin expansion in ascending powers of x, we may derive the value of

$$\left\{\frac{d^{n-s}}{dx^{n-s}}\left[\frac{dz}{dx}\{\phi(z)\}^{s-1}F_{s-1}(0,z)\right]\right\}_{x=0}$$

from the Lagrangian expansion of

$$\frac{dz}{dx}\{\phi(z)\}^{s-1}F_{s-1}(0,z)$$

in ascending powers of x. The coefficients of $x^n/n!$ is thus seen to have the form

$$\int_0^a F_n[a-\tau,\tau][\phi(\tau)]^n d\tau+F_{n-1}[0,a]\,[\phi(a)]^n+\frac{d}{da}\{F_{n-2}[0,a][\phi(a)]^n\}$$
$$+\cdots+\frac{d^{n-1}}{da^{n-1}}\{F[0,a][\phi(a)]^n\}\ .$$

An editor has kindly remarked that our expansion problem can be regarded as a particular case of the following more general expansion problem.

Let $w=f(x, z)$ be a function which is holomorphic in the neighborhood of $x=0$ and $z=a$, let $\phi(z)$ be holomorphic in the neighborhood of $z=a$ and let $z=z(x)$ be that solution of the equation $z=a+x\phi(z)$ which is holomorphic in the neighborhood of $x=0$ and tends to the value a as $x\to 0$. The problem is to expand w in a power series in x convergent for sufficiently small values of $|x|$.

This problem, in its turn, is merely a special case of a more general problem treated by Cauchy.*

Given a function $g(z, x)$ holomorphic for $|z|<r$, $|x|<\rho$, for which $z=0$ is an m-fold 0 of $g(z, 0)$. Let $r_1<r$ and $\rho_1<\rho$ be so chosen that for $|x|\leqq\rho_1$, the function $g(z, x)$ is different from zero on the circle $|z|=r_1$ and admits of m zeros z_1, z_2, z_m inside of this circle, which zeros are continuous functions of x, vanishing with x. Finally let $f(x, z)$ be an analytic function of x and z which is holomorphic for $|x|<\rho_1$, $|z|<r_1$. Then the sum

$$F(x)=f(x,z_1)+f(x,z_2)+\cdots+f(x,z_m)$$

is a holomorphic function of x in $|x|<\rho_1$ and expressions can be found for the coefficients in the Maclaurin series for $F(x)$ in powers of x.

3. **Special cases of the expansion.** When $a=0$, $\phi(\tau)=1$, $F=F(x-\tau, \tau)$ we have the expansion

$$\int_0^x F[x-\tau,\tau]\,d\tau=\frac{x}{1!}F[0,0]+\frac{x^2}{2!}\Big[F_{10}[0,0]+F_{01}[0,0]\Big]$$
$$+\frac{x^3}{3!}\Big[F_{20}[0,0]+F_{11}[0,0]+F_{02}[0,0]\Big]+\cdots,$$

where now

$$F_{mn}[\sigma,\tau]=\frac{\partial^{m+n}}{\partial\sigma^m\partial\tau^n}\,F[\sigma,\tau]\ .$$

* E. Lindelöf, *Calcul des Résidus*, 1905, p. 24.

In particular, if $F=f(x-\tau)g(\tau)$ where $f(x)$ and $g(x)$ are analytic in the neighborhood of $x=0$, we have the result that for sufficiently small values of x, the definite integral

$$h(x)=\int_0^x f(x-t)g(t)dt$$

can be expanded in the form

$$h(x)=\frac{x}{1!}f(0)g(0)+\frac{x^2}{2!}[f'(0)g(0)+f(0)g'(0)]$$
$$+\frac{x^3}{3!}[f''(0)g(0)+f'(0)g'(0)+f(0)g''(0)]+\cdots,$$

where primes denote differentiations with respect to the argument. Thit result occurs in a slightly different form in Borel's theory of divergens series.* Writing

$$f(x)=f(0)+\frac{x}{1!}f'(0)+\frac{x^2}{2!}f''(0)+\cdots,$$
$$g(x)=g(0)+\frac{x}{1!}g'(0)+\frac{x^2}{2!}g''(0)+\cdots,$$
$$F(y)=\int_0^\infty e^{-xy}f(x)dx,\quad G(y)=\int_0^\infty e^{-xy}g(x)dx,\quad F(y)G(y)=\int_0^\infty e^{-xy}h(x)dx,$$
$$F(y)=f(0)y^{-1}+f'(0)y^{-2}+f''(0)y^{-3}+\cdots,$$
$$G(y)=g(0)y^{-1}+g'(0)y^{-2}+g''(0)y^{-3}+\cdots,$$
$$F(y)G(y)=f(0)g(0)y^{-2}+[f'(0)g(0)+f(0)g'(0)]y^{-3}+\cdots,$$

the form of the series for $h(x)$ is at once suggested.

Since integral equations of type

$$h(x)=\int_0^x f(x-\tau)\,g(\tau)d\tau$$

are of some importance in analysis† it should be worth while to consider the more general class of integral equations of type

* *Leçons sur les Séries Divergentes*, Paris, 1901. This expansion and the previous one may be obtained by writing $\tau=sx$ in the integral and expanding in powers of x by Maclaurin's theorem. The most direct method, however, is that given at the end of §2.

† A method of solving equations of this type which depends on a determination of the functions $F(y)$ and $G(y)$ was suggested by Vilfredo Pareto, Journal für Mathematik, vol. 110 (1892), p. 290. The method was suggested again in a more general form by the present author *Report on integral equations*, British Association Report, 1910, and was actually used in the solution of

$$h(x)=\int_a^z F[a+x\phi(\tau)-\tau]g(\tau)d\tau\ ,$$

where

$$z=a+x\phi(z)\ ,$$

and $g(\tau)$ is the function to be determined.

When the definite integral can be expanded in a convergent power series

$$h(x)=\frac{x}{1!}F(0)g(a)\phi(a)+\frac{x^2}{2!}\Big[F'(0)\{g(a)\}\{\phi(a)\}^2$$
$$+F(0)\frac{d}{da}\{g(a)[\phi(a)]^2\}\Big]+\frac{x^3}{3!}\Big[F''(0)g(a)\{\phi(a)\}^3$$
$$+F'(0)\frac{d}{da}\{g(a)[\phi(a)]^3\}+F(0)\frac{d^2}{da^2}\{g(a)[\phi(a)]^3\}\Big]+\cdots\ ,$$

the quantities $g(a)$, $g'(a)$, $g''(a)$, $\cdots$ can be determined uniquely from the coefficients in the known power series for $h(x)$. The function $g(\tau)$ can then be calculated for other values of τ by Taylor's theorem. This solution is, of course, of a purely formal nature and must be supplemented by convergence theorems.

4. **An application of Lagrange's expansion.** To prove that the quantity θ of § 1 is a solution of

$$f\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\theta=0\ ,$$

it is convenient to establish the formula

$$g\left(\frac{\partial}{\partial x},\frac{\partial}{\partial y},\frac{\partial}{\partial z}\right)\theta=\left(\frac{1}{M}\frac{\partial}{\partial\theta}\right)^{m-1}\left\{\frac{g[\xi(\theta)\ ,\eta(\theta)\ ,\zeta(\theta)]}{M}\right\}, \tag{1}$$

where $g(\xi,\eta,\zeta)$ is any homogeneous polynomial of degree m in ξ, η, ζ. Consider the equation

$$(x+a)\xi(\omega)+(y+b)\eta(\omega)+(z+c)\zeta(\omega)=\chi(\omega)\ .$$

Writing

$$\sigma=\chi(\omega)-x\xi(\omega)-y\eta(\omega)-x\zeta(\omega)\ ,$$

$$\omega=\phi(\sigma)\ ,\qquad \theta=\phi(0)\ ,$$

$$\sigma=a\xi[\phi(\sigma)]+b\eta[\phi(\sigma)]+c\zeta[\phi(\sigma)]\ ,$$

---

a problem in Messenger of Mathematics, vol. 49 (1920), p. 1. The method has been used recently by J. R. Carson in a number of articles appearing in the Bell System Technical Journal.

and introducing a parameter t which is afterwards put equal to unity we may write the last equation in the form

$$\sigma = t\{a\xi[\phi(\sigma)] + b\eta[\phi(\sigma)] + c\zeta[\phi(\sigma)]\} .$$

We may now expand the function $\omega = \phi(\sigma)$ in ascending powers of t by Lagrange's theorem* if a, b and c are sufficiently small. The expansion is

$$\omega = \phi(\sigma) = \phi(0) + \sum_{n=1}^{\infty} \frac{t^n}{n!} \left\{ \frac{d^{n-1}}{d\sigma^{n-1}} \left[\phi'(\sigma) \{a\xi[\phi(\sigma)] + b\eta[\phi(\sigma)] + c\zeta[\phi(\sigma)]\} \right] \right\}_{\sigma=0}.$$

Putting $t=1$ and transforming back to the variable θ we obtain the expansion

$$(2) \qquad \omega = \theta + \sum_{n=1}^{\infty} \frac{1}{n!} \left(\frac{1}{M} \frac{\partial}{\partial\theta} \right)^{n-1} \left[\frac{\{a\xi(\theta) + b\eta(\theta) + c\zeta(\theta)\}^n}{M} \right] .$$

Regarding this as the Taylor expansion of ω in powers of a, b and c, we may calculate the different partial derivatives of order m with respect to the three independent variables x, y and z by finding the coefficients of the different products of powers of a, b and c. The general result can be expressed in the form (1). The result of § 1 may be derived from that of § 2 with the aid of the above device of introducing an auxiliary variable t.

This device is a familiar one in the theory of Taylor series in several variables and the theorem used in obtaining (2) is really one generalization of Lagrange's theorem for the case of several independent variables. An entirely different generalization giving a power series in more than one variable has been obtained by Darboux† for the case of two variables and by Stieltjes‡ for the case of n variables. Stieltjes remarks that the method by which Heine derived Lagrange's expansion by the calculus of variations§ can be generalized so as to give more general expansions.|| Many other generalizations of Lagrange's expansion are known. Besides

* J. L. Lagrange, Berlin Memoirs, 1768.

† G. Darboux, Comptes Rendus, vol. 68, p. 324; C. Hermite, *Cours d'Analyse*, 4th edition, p. 182; H. Poincaré, Acta Mathematica, vol. 9.

‡ T. J. Stieltjes, Annales de l'Ecole Normale, ser. 3, vol. 2 (1885), p. 93. Collected works, vol. 1, p. 445.

§ E. Heine, Journal für die reine und Angewandte Mathematik, vol. 54 (1857).

|| An application of the general expansion of Stieltjes is mentioned in a paper by the author, Bulletin of the American Mathematical Society, vol. 22 (1916), p. 329. There is, however, a mistake in sign in equations (10). These equations should read $\sigma_p = \mu_p - \xi_p\theta$, $\tau_p = v_p - \eta_p/\theta$.

those due to Bürmann,* Teixeira† and Rouché‡ that are given in Whittaker and Watson's *Modern Analysis*, there is a recent generalization due to Kössler.§ References to some further developments are given below||.

---

* Bürmann, Mémoires de l'Institut, vol. 2 (1799), p. 13. See also A. C. Dixon Proceedings of the London Mathematical Society, ser. 1, vol. 34 (1902) p. 151.

† F. G. Teixeira, Resal Journal, ser. 3, vol. 7 (1881), p. 277. Académie Royale de Belgique, Mémoires in 4 to, ser. 2, vol. 1. Journal für die reine und angewandte Mathematik, vol. 122 (1900), p. 97.

‡ E. Rouché, Paris Memoirs, vol. 18 (1868), p. 422.

§ M. Kössler, Proceedings of the London Mathematical Society, ser. 2, vol. 20 (1922), p. 365.

|| G. Zolotareff, Nouvelles Annales, ser. 2, vol. 15 (1876), p. 422.
E. McClintock, American Journal of Mathematics, vol. 4 (1881), p. 16.
L. W. Meech, The Analyst, vol. 3, p. 33.
N. W. Bugajew, Moscow Mathematical Papers, vol. 22 (1901), p. 219.
M. Lerch, Rozpravy, vol. 29 (1911), p. 14.
A. R. Johnson, Messenger of Mathematics, ser. 2, vol. 14 (1885), p. 76.

California Institute of Technology,
Pasadena, Calif.